Efforts to address the problems of distressed urban neighborhoods stretch back to the 1800s, but until relatively recently, data played little role in forming policy. It wasn't until the early 1990s that all of the factors necessary for rigorous, multifaceted analysis of neighborhood conditions—automated government records, geospatial information systems, and local organizations that could leverage both—converged. Strengthening Communities documents that convergence and details its progress, plotting the ways data are improving local governance in America.

For organizers, advocates, civic leaders, local government officials, and policy researchers who want to understand how data can advance community change, this book is a huge gift. With an eye on future challenges, the authors lay the groundwork for animating democracy by demonstrating how effective use of data can deepen participation. This book could have been called "Data Rock!"

— Angela Glover Blackwell
Founder and CEO, PolicyLink

The 21st century has brought vast amounts of new data on conditions and trends in urban communities. We are only beginning to grasp the value and power of this information. This book should be required reading for researchers who want to harness new sources of data to gain a deeper understanding of how and why neighborhoods change. It is also a useful roadmap for community leaders and public officials who want to use data to identify emerging challenges in low-income neighborhoods and craft effective policies to address them.

— Dr. Ingrid Gould Ellen
Professor and Program Director, NYU

The sea change we have seen in the collection and use of neighborhood data represents a true advancement in our ability to conduct evidence-based policymaking. This engaging book tells the story of this revolution and offers a vital roadmap for those seeking to contribute to lasting improvements in America's communities. It is a must read for anyone aspiring to be a positive change agent at the neighborhood level, especially in neighborhoods that are home to low-income and minority families.

— Dr. Raphael W. Bostic
Bedrosian Center Director, USC Price School of Public Policy, and former HUD Assistant Secretary for Policy Development and Research

Data are among the most powerful tools available in a democracy. Armed with data, communities can cut through ideological boundaries, focus on things that matter, and engage in conversations about challenges and opportunities. Anyone interested in helping a community make good decisions should read this book to understand the use of neighborhood data and the community information field.

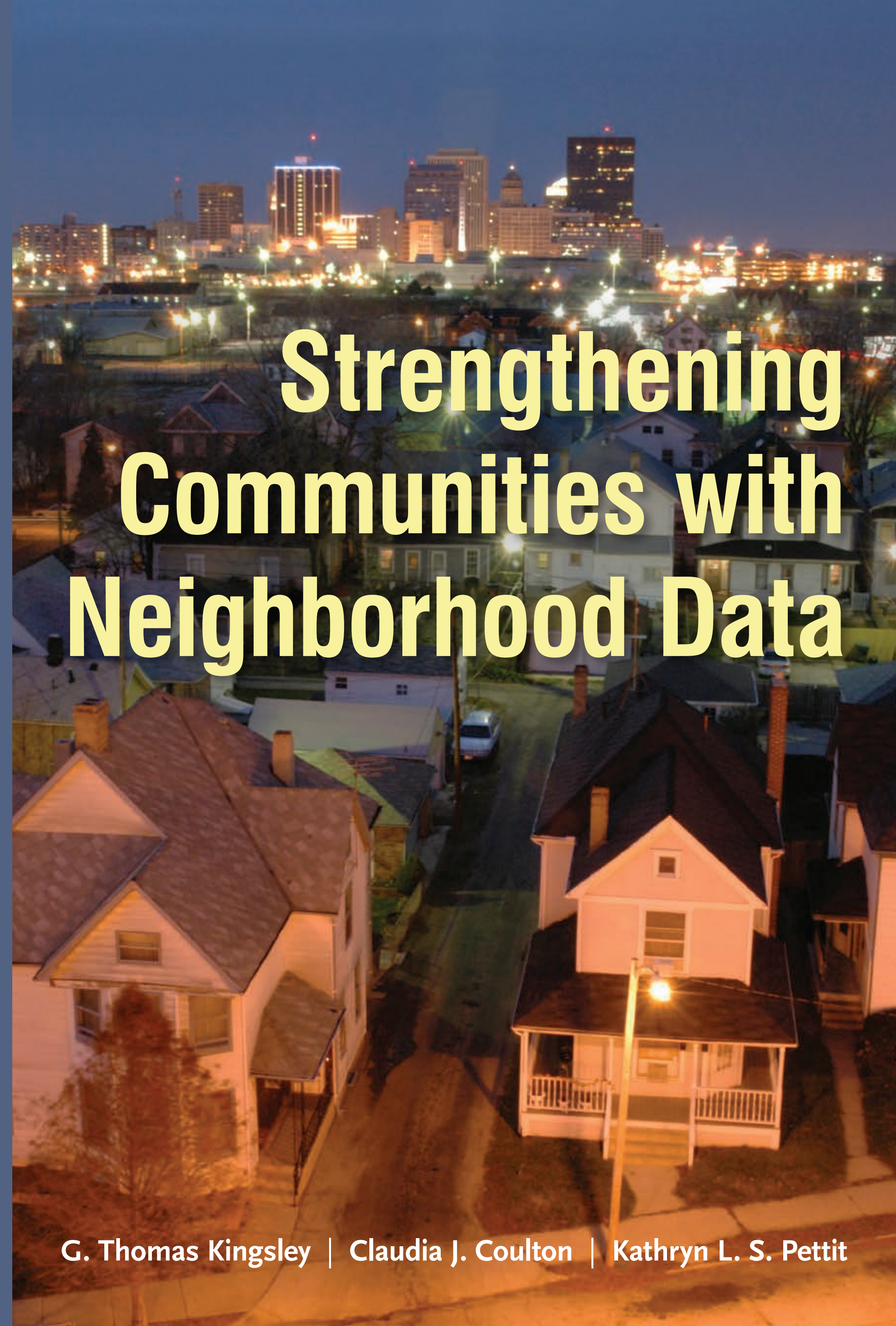
— Paul S. Grogan
President and CEO, The Boston Foundation

*Visit* **http:/www.urban.org/StrengtheningCommunities**
*to download the free ebook or order a hard copy.*

Strengthening Communities with Neighborhood Data

G. Thomas Kingsley
Claudia J. Coulton
Kathryn L. S. Pettit

URBAN INSTITUTE



# Strengthening Communities with Neighborhood Data

G. Thomas Kingsley

Claudia J. Coulton

Kathryn L. S. Pettit

URBAN INSTITUTE · ELEVATE · THE · DEBATE

**ABOUT THE URBAN INSTITUTE**

The nonprofit Urban Institute is dedicated to elevating the debate on social and economic policy. For nearly five decades, Urban scholars have conducted research and offered evidence-based solutions that improve lives and strengthen communities across a rapidly urbanizing world. Their objective research helps expand opportunities for all, reduce hardship among the most vulnerable, and strengthen the effectiveness of the public sector.



URBAN
INSTITUTE · ELEVATE · THE · DEBATE

# Contents

# Acknowledgments

We would like first to acknowledge Craig Howard of the John D. and Catherine T. MacArthur Foundation for his recognition of the need to document the state of the community information field and his support of our efforts to do so. We also would like to thank the contributing essayists, whose knowledge and experience greatly enhanced the book.

Many others supported us along the way in the completion of this volume. Leah Hendey, James Gibson, and Michael Rich provided very helpful feedback on early drafts of the chapters. We also had the assistance of several colleagues who reviewed individual sections: Jennifer Comey, Susan Millea, Taryn Roch, Aaron Truchil, Susana Vasquez, and Chris Walker. We recognize two leaders who facilitated the Urban Institute's role in the National Neighborhood Indicators Partnership (NNIP): James Gibson, who while at the Rockefeller Foundation funded the local work that was the basis for the NNIP concept, and William Gorham, the Urban Institute's founding president, who understood how NNIP could advance the Institute's mission. In addition, the Annie E. Casey Foundation, represented by Cindy Guy, provided extraordinary financial and intellectual support to help NNIP grow and to further the use of data to improve conditions for children and families. Alaina Harkness from the John D. and Catherine T. MacArthur Foundation has contributed invaluable encouragement to the NNIP network and broader leadership related to philanthropy's role in building local capacity to use data for community action.

Finally, we thank the staff members of the local NNIP organizations and other indicator practitioners, who work tirelessly to make data available and understandable so that their communities can make more informed and equitable decisions.

This book was supported by the John D. and Catherine T. MacArthur Foundation. The views expressed are those of the authors and do not necessarily represent those of the foundation or the Urban Institute, its trustees, or its funders.

*G. Thomas Kingsley*
*Claudia J. Coulton*
*Kathryn L. S. Pettit*

# 1

# Introduction to the Field

## The Inception of a Field

America's cities have never been homogenous. Most have always had neighborhoods of considerable wealth, with broad tree-lined streets and spacious houses. American cities have also had poorer neighborhoods that range from physically run-down areas to deeply troubled, crime-ridden, and congested slums. And between these disparate types of neighborhoods, they have had many middle- and working-class areas.

Efforts to address the problems of distressed neighborhoods stretch back to the 1800s, but they did not gain much prominence nationally until the emergence of the settlement house movement and social reforms of the early 20th century. A variety of approaches have been tried since, often with contrasting philosophical underpinnings: urban renewal, model cities, community organizing, community development.[1] There have been many notable success stories along the way, but a not infrequently heard conclusion is that these approaches generally fell short of their designers' aspirations (see, for example, Kubisch et al. 2010). A century later, the problems of distressed neighborhoods are still very much with us, heightened by the foreclosure crisis and Great Recession of the past decade.

Before 1990 at least, none of these policy approaches were much guided by data. Why was that the case? Relevant statistics (e.g., rates of crime, teen pregnancy, welfare recipiency, and property tax delinquency) were often available for cities as a whole, but almost never for individual neighborhoods. This lack of neighborhood-level data resulted from the prohibitive expense of plotting the locations of thousands of transactions that backed up each of these indicators and then summarizing the results by neighborhood.

Some researchers and practitioners saw the lack of data as a serious problem early on. A good example is the view expressed by Ahlbrandt and Brophy in their 1975 book *Neighborhood Revitalization*:

> The formulation of a strategy to stem decline in an urban neighborhood requires an understanding of the specific conditions that prevail in that location. This necessitates the collection and evaluation of data describing the overall direction of change in addition to the specific conditions affecting the viability of the neighborhood housing market. . . . The measurement of housing and neighborhood conditions has traditionally been expressed in physical terms . . . [that] fall short of gauging the overall neighborhood conditions that comprise the larger housing environment. . . . This larger neighborhood environment can be measured by looking at the economic, social, psychological and demographic aspects of the area. (53)

Some small-area data were available from the US Census Bureau's decennial censuses, but these did not include many indicators that were of value in neighborhood planning, and updates of those few that were available were frustratingly infrequent.

This dearth of neighborhood data was not just a problem for revitalization planners. Police departments knew they needed to understand many social, economic, and physical aspects of neighborhoods to spatially target crime-fighting resources efficiently. Foundations realized they could hardly begin to reliably understand what their grants were accomplishing without factual information on how neighborhoods differed from each other and how they were changing.

But in the early 1990s, a transformative new possibility emerged as two trends reached critical thresholds. First, more and more local government agencies had automated their administrative records, so that records on transactions that contained an address or some other geographic identifier were now computer based. Second, geographic information system technology had reached a point at which it could process, plot, and manipulate such data with considerable efficiency.

The payoff from these two developments, however, required yet another innovation. Some entity in a city had to make it a part of its mission to collect these data from various local agencies, enter the data into an orderly information system, make the data available for use, and keep the data up to date. The first entity to fulfill this commitment was a community-oriented university research center: the Center on Urban Poverty and Social Change at Case Western University in Cleveland, Ohio, in 1992. For the first time, it became possible to track changing neighborhood conditions annually, between censuses, by using a variety of indicators (a development first documented by Coulton in 1995).

As the work got under way, this university-based center was part of the Community Planning and Action Program, a small network of local groups funded by the Rockefeller Foundation to develop local solutions to emerging concerns about concentrated urban poverty in several US cities. Three of these groups (from Boston, Massachusetts; Denver, Colorado; and Oakland, California), inspired by the neighborhood indicators idea and the progress being made in Cleveland, decided to build similar systems. Members of the Community Planning and Action Program then became aware of similar developments in Atlanta, Georgia, and Providence, Rhode Island, and met with their representatives to explore ideas about how the approach might be strengthened and spread to other places.[2]

They decided it would make sense to form a new network of organizations doing this work, but that to be sustainable the network should be based in an established public policy research institute. They asked the Urban Institute if it would be interested in playing that role. The Urban Institute, a nonprofit research organization that seeks to foster sound public policy and effective government, seemed a good fit for the organization's mission. After a period of study, the Institute agreed, and they jointly formed the National Neighborhood Indicators Partnership (NNIP) in early 1996. The Institute study of what the partners were accomplishing yielded the partnership's first guidebook, *Building and Operating Neighborhood Indicator Systems* (Kingsley 1999).

Although NNIP partner organizations used their data to address policy and programmatic issues at various levels, a primary motivation of the partnership was to bring data to bear to improve conditions in distressed low-income neighborhoods. Their theme was "democratizing information," implying a commitment to directly engage neighborhood practitioners and residents in using data in community improvement efforts.[3]

## The Purpose and Focus of This Book

In the two decades since the first NNIP data system was developed, dramatic changes have occurred in the community information field and its context. First, the partnership itself has expanded, with local partners now in about three dozen cities. The local partners and the partnership have completed a substantial body of work, including policy and community-building applications, technical innovations, and guidebooks for practitioners.[4]

But the nature of the work of the local partner organizations and the environment in which they operate have also been markedly affected by other important developments:

- Technological advances have yielded unprecedented reductions in the costs of data assembly, storage, manipulation, and display.
- The amount of relevant data available to the public has been vastly expanded. These sources include new national data files with small-area or address-level data (from the federal government and commercial sources), as well as publicly available local government administrative files and data available from commercial sources.
- Data visualization platforms and online tools have been developed that make it easier for users to work with neighborhood-level data.
- Many local governments have markedly improved their own internal data capacities (e.g., staff knowledge, data collection, program and policy applications).
- More outside consultants are now available who can help local organizations take advantage of these new capacities.

These changes have already enabled a wide variety of users to apply data more effectively, not only in practical efforts to address short-term issues, but also in research that furthers understanding of the process of community change, thereby setting the stage for better solutions in the long term. To date, however, documentation of these changes and their impacts has been partial and fragmentary.

The purpose of this book is to rectify that deficit—to tell the story of how the community information field has evolved over the past two decades in a balanced manner and to identify ways it has influenced community change. Our aim is not only to document the advances, but to set them in a theoretical context in a manner that will serve as a sound basis for guidance on how the field should evolve.

The story is not just about data. It is also about how changes in a host of factors (institutional, technical, political, philosophical) have come together and interacted with expanded data availability to alter the way governance at the local level is being conducted in America. And the story is not just about NNIP. The three authors of this book have worked actively with NNIP for many years, and we think the NNIP approach—that is, integrating and making available a variety of data for small sub-city areas and locations, with primary emphasis on improving conditions in low-income neighborhoods—has been critical to the development of the field. But in this book, we also document the work of other actors who have made key contributions to the field.

Rather than trying to examine all implications, the book concentrates on two particularly important areas. First, we focus on the development and use of data at the local level in metropolitan areas. National actors both contribute and use neighborhood-level data, but work at the local level generates what are by far the dominant payoffs. Because NNIP partners are all in metropolitan areas, the local experiences described in this book all occur in urban contexts. Many of the lessons and concepts are relevant to small towns or rural counties, but others may need adaptation for less-populated areas. Second, we focus on applications designed to achieve collective ends mostly related to governance broadly defined; applications are intended for use by institutions rather than individuals. Many valuable applications of the newly available data are used directly by individuals (e.g., real-time bus schedules downloaded on cell phones and websites that help families find and take advantage of services they need). Such applications are important and warrant documentation, but they are not covered in this book.

## Key Concepts

In this section, we present our definitions of two topics that frame the remainder of the book: community and indicators.

### What Is Community?

In this book we focus on place-based communities. A large body of scientific evidence indicates that where people live matters for their well-being (Ellen and Turner 1997). The quality of local public services, the prevalence of crime and violence, the quality of the natural and built environments, the influences of peers and social networks, and proximity to jobs and resources can all act either to diminish the well-being of individuals or enhance their prospects. A substantial body of research finds that growing up in disinvested, distressed, or socially and economically isolated neighborhoods is associated with an increased risk of many adverse outcomes for children, including school failure, poor health, victimization, delinquency, teen childbearing, and youth unemployment (Brooks-Gunn, Duncan, and Aber 1997; Ellen, Mijanovich, and Dillman 2001; Leventhal and Brooks-Gunn 2003). The influence of the places where these children live persists throughout their life course, especially as there is a strong chance, despite residential mobility, that those who grow up in distressed areas live in similar areas as adults (Sharkey 2008). Nevertheless, when families are assisted in moving out of distressed areas and into middle-class neighborhoods, individuals can experience improvements in their subjective well-being and many aspects of health (Orr et al. 2003; Ludwig et al. 2008; Keels et al. 2005; Sanbonmatsu et al. 2012), suggesting the importance of neighborhood context along the continuum of advantage.

In this book, we use the terms *neighborhood* and *place-based community* interchangeably. Although it is important to recognize that communities can be far-flung and even virtual, this book focuses on geographically defined areas that provide an important context for the well-being of individuals and the success of the regions in which they are located. Typically referred to as neighborhoods within urban areas, they can also comprise villages or hamlets in rural areas. These areas are not simply geographic footprints but units of social organization that have meaning as places to live, work, and go about daily life. They have an identity in the minds of insiders and outsiders. Neighborhoods and villages are more than collections of individuals or locations for populations; they also include space, physical structures, social networks, formal and informal organizations, businesses, systems of exchange and governance, and so forth. These place-based communities are not islands, but are spatially located relative to other places. Moreover, they operate at various scales, from the immediate residential vicinity to wider areas of social, economic, and political relevance to daily life.

The problem of defining neighborhoods and the practical issue of specifying their boundaries for local indicators work is taken up in detail in chapter 7. However, it is important to recognize throughout this volume that the concept of neighborhood is multidimensional and dynamic (Chaskin 1997; Galster 2001; Nicotera 2007) and that the delineation of place-based communities requires careful consideration of the particular purpose of the analysis. Indeed, although historically neighborhood indicators work has relied on statistical definitions driven by census geography or boundaries drawn by government agencies or planning groups, increasingly analysts are using flexible and overlapping geographic units and incorporating multiple perspectives into delineating place-based communities that are relevant to the particular issue at hand. Clearly, an appreciation of local knowledge and of the fact that areas are often contested is a crucial underpinning of neighborhood indicators work.

## Community Indicators and Measures

In this book we will often refer to community indicators. These are measures of living conditions, resources, or attributes of populations that can be compared over time or between places and groups. Additionally, the term *indicator* implies that the measure is a statistic that is being used for assessment purposes (Bradburn and Fuqua 2010). Indicators across many domains are required to fully assess the overall well-being of a society. Key dimensions include material living standards, shelter, health, education, personal activities, political voice, social connectedness, environment, and security (Stiglitz, Sen, and Fitoussi 2009). For each of these dimensions, numerous specific measures have been proposed, but in practice the choice of measures is often dictated by the availability of data series that are uniform across time and space. To be used as indicators, the measures have to be precise enough to detect differences over time and among communities. Moreover, because community indicators are designed to support action, the measures should possess enough face validity to be believable and sufficient construct validity to be interpretable. The various roles indicators can play in community work and local governance are more broadly discussed in chapter 4. Methodological issues in developing measures for community indicators are covered in chapter 7.

Localized community indicators are necessary because there is considerable variation in human well-being depending on where individuals reside within a city or region. Indicators for the city as a whole or larger geographies mask these differences. Indicators in selected locations can reveal groups of individuals in great distress even when things are generally improving for the population overall. Indeed, the existence of inequality of human well-being is often starkly revealed when indicators in one community are compared with another or with national or statewide averages. Such differences often provide justification for changes in public policy, program delivery, or distribution of resources. Disparities in human well-being can guide efforts to mobilize communities to act on improving conditions for themselves. And local indicators that reveal pockets of concern about particular groups can be used to target resources to areas where they are needed most.

## The Structure and Audience of This Book

The book's chapters are written by the authors. Most chapters are followed by one or more essays that have been solicited from experts in the field. The essays are not commentaries on the chapters but, rather, provide original material reflecting the experience and thinking of the contributors.

- Chapter 2 tells the institutional side of the story. It first describes the roles being played by local institutions and how they have changed as the information revolution has unfolded. It then describes the NNIP partner organizations: who they are, what they do, and how they do it.
- The data and technical sides are documented in chapter 3. The chapter begins by reviewing relevant data that are now available for community work and the way the task of data assembly has changed since 1996 (including the role of the open data movement). It then examines the main advances in technology that are transforming practice in all aspects of the field.
- Chapter 4 introduces a framework for considering how community information can be applied. It first describes the multiple uses of indicators foreseen when the social indicators movement began, and then discusses five basic types of applications that illustrate how community information can be made useful in local decision-making today.
- The next two chapters review how the newly available administrative data are being applied to achieve practical ends in efforts to

improve and provide services to individual neighborhoods (chapter 5) and to develop strategies for neighborhoods across a city or region (chapter 6).

- Chapter 7 focuses on methods of analysis. It offers a framework for identifying methodologies pertinent to neighborhood indicators and reviews selected techniques and tools that represent promising approaches to addressing the range of applications for which neighborhood and community data can be employed.
- Finally, in chapter 8, the authors consider the implications of the earlier chapters and suggest future directions for the field and recommendations to enhance its development.

This book has been written to be of value to several audiences. One is professionals and practitioners who are interested in developing data intermediary capacities in their own cities since the cities that currently have NNIP partners represent only a small fraction of all urban areas in the nation. Equally important, as interest grows in data-driven decisionmaking and accountability, many others at the local level (civic leaders, grass roots community groups, local government staffs) need to understand the state of the art and the possibilities of this field. In addition to these two audiences, the number of researchers interested in doing work related to community change is growing. This book should provide important guidance to their efforts by offering them a solid baseline understanding of what has been done, thus helping them to avoid redundancy and to focus their own work more efficiently. We also believe this book will prove valuable to students pursuing careers in local policy, planning, and community building and in urban research and to the educators who guide them in the quest for relevant knowledge. Finally, we seek to influence the national actors who support the use of neighborhood-level data in local decisionmaking. We hope that raising awareness of the community information field for practice and research among local and national stakeholders will ultimately result in more informed and inclusive decisionmaking in communities across the nation.

## NOTES

1. Reviews of these programmatic approaches are provided by Von Hoffman (2012) and Rohe (2009).

2. James O. Gibson, a Rockefeller Foundation Program Manager, mobilized and directed the Community Planning and Action Program. The Cleveland system was built

by Claudia J. Coulton, Director of the Center on Urban Poverty and Social Change, and her staff. The other three organizations (and their directors) were the Boston Foundation Persistent Poverty Project (Charlotte Kahn); the Piton Foundation in Denver (Terri J. Bailey); and the Urban Strategies Council in Oakland, California (Angela Blackwell, assisted by Joaquin Herranz). The groups in the two other cities were the Atlanta Project (David Sawicki) and The Providence Plan (Michael Rich and Pat McGuigan). An independent consultant, Talton Ray, also played a critical role in bringing these groups together.

3. Sawicki and Craig (1996) applied this term as a central NNIP principle and provide a longer exploration of this topic.

4. NNIP has been supported over the years by several national foundations. The Annie E. Casey Foundation, with Cynthia Guy as grant manager, has provided funding consistently since the partnership was formed. Other major funders have included the John D. and Catherine T. MacArthur Foundation (Craig Howard and Alaina Harkness) and in the early years, the Rockefeller Foundation (Darren Walker). Initial funders also included the James Irvine Foundation and the Surdna Foundation.

## REFERENCES

Ahlbrandt, Roger S., Jr., and Paul C. Brophy. 1975. *Neighborhood Revitalization.* Lexington, MA: Lexington Books.

Bradburn, N. M., and C. J. E. Fuqua. 2010. "Indicators and the Federal Statistical System: An Essential but Fraught Partnership." *Annals of the American Academy of Political and Social Science* 631: 89–108.

Brooks-Gunn, J., G. J. Duncan, and J. L. Aber. 1997. *Neighborhood Poverty: Context and Consequences for Children.* New York: Russell Sage Foundation Publications.

Chaskin, R. J. 1997. "Perspectives on Neighborhood and Community: A Review of the Literature." *Social Service Review* 71 (4): 521–47.

Coulton, Claudia J. 1995. "Using Community Level Indicators of Children's Well-Being in Comprehensive Community Initiatives." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Context*, edited by James P. Connel, Anne C. Kubisch, Lisbeth Schorr, and Carol H. Weiss. Washington, DC: The Aspen Institute.

Ellen, I. G., T. Mijanovich, and K. Dillman. 2001. "Neighborhood Effects on Health: Exploring the Links and Assessing the Evidence." *Journal of Urban Affairs* 23 (3–4): 391–408.

Ellen, I. G., and M. A. Turner. 1997. "Does Neighborhood Matter: Assessing Recent Evidence." *Housing Policy Debate* 8 (4): 833–66.

Galster, G. 2001. "On the Nature of Neighbourhood." *Urban Studies* 38 (12): 2111–124.

Keels, M., G. J. Duncan, S. DeLuca, R. Mendenhall, and J. Rosenbaum. 2005. "Fifteen Years Later: Can Residential Mobility Programs Provide a Long-Term Escape from Neighborhood Segregation, Crime, and Poverty?" *Demography* 42 (1): 51–73.

Kingsley, G. Thomas, ed. 1999. *Building and Operating Neighborhood Indicator Systems.* Washington, DC: The Urban Institute.

Kingsley, G. Thomas, Claudia J. Coulton, Michael Barndt, David S. Sawicki, and Peter Tatian. 1997. *Mapping Your Community: Using Geographic Information to*

*Strengthen Community Initiatives.* Washington, DC: US Department of Housing and Urban Development.

Kubisch, Anne C., Patricia Auspos, Prudence Brown, and Tom Dewar. 2010. *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts.* Washington, DC: The Aspen Institute.

Leventhal, T., and J. Brooks-Gunn. 2003. "Children and Youth in Neighborhood Contexts." *Current Directions in Psychological Science* 12 (1): 27–31.

Ludwig, J., J. R. Kling, L. F. Katz, L. Sanbonmatsu, J. B. Liebman, G. J. Duncan, et al. 2008. "What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment?" *American Journal of Sociology* 114 (1): 144–88.

Nicotera, N. 2007. "Measuring Neighborhood: A Conundrum for Human Services Researchers and Practitioners." *American Journal of Community Psychology* 40 (1–2): 26–51.

Orr, L., J. D. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, L. F. Katz, et al. 2003. *Moving to Opportunity for Fair-Housing Demonstration: Interim Impacts Evaluation.* Washington, DC: US Department of Housing and Urban Development.

Rohe, William M. 2009. "From Local to Global: One Hundred Years of Neighborhood Planning." *Journal of the American Planning Association* 75 (2): 209–30.

Sawicki, David S., and William J. Craig. 1996. "The Democratization of Data: Bridging the Gap for Community Groups." *Journal of the American Planning Association* 62 (4): 512–23.

Sanbonmatsu, L., N. A. Potter, E. Adam, G. J. Duncan, L. F. Katz, R. C. Kessler, T. W. McDade, et al. 2012. "The Long-Term Effects of Moving to Opportunity on Adult Health and Economic Self-Sufficiency." *Cityscape* 14 (2, Moving To Opportunity): 109–36.

Sharkey, P. 2008. The Intergenerational Transmission of Context. *American Journal of Sociology* 113 (4): 931–69.

Stiglitz, J., A. Sen, and J. E. Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress.* http://www.citymaking.com/wp-content/uploads/2010/01/19784660-Happiness-and-Measuring-Economic-Progress-by-Joseph-Stiglitz.pdf.

Von Hoffman, Alexander. 2012. "The Past, Present, and Future of Community Development in the United States." In *Investing in What Works for America's Communities: Essays on People, Places and Purpose*, edited by Nancy O. Andrews and David J. Erickson. San Francisco: Federal Reserve Bank of San Francisco and Low-Income Investment Fund.

# 2
# Institutional Context

The local institutional context has a great influence on how community information will be developed and applied in an area, but the pattern of influence is a two-way street. On one hand, local institutions have missions that drive their basic demands for information. External forces (such as the shifting priorities given to different policy issues, the changing fiscal environment, and advances in programmatic best practices) modify not only the nature of the activities they carry out but also the definition of their missions. These shifting emphases change the information they need and the way they use it. These changes, in turn, put pressure on the institutions that produce information to shift their own agendas (note that the prime local users of community information are often prime producers of the original data).

On the other hand, the broader availability of information and the new technology influence the institutional context. They can lead to the identification of new problems and opportunities and improve understanding of the efficacy of different policies and programs. That ability, in turn, can cause user institutions to alter their mission statements as well as their programmatic agendas, thus altering their further demands for and use of information. In short, it is necessary to understand the dynamic institutional context in order to understand both recent and future directions for this field.

The two major sections of this chapter explore these relationships and how they have changed over the past two decades. The first section discusses local institutions and community information. It identifies the types of individual institutions that have information-related functions. It then examines how these institutions come together to make decisions about policies and programs to achieve broader collective goals. The section closes with a discussion of the roles of state- and federal-level actors in influencing local information development and use.

The second major section discusses local data intermediaries. Despite considerable advances, most communities still face serious barriers to the effective development and use of community information, barriers that arise mostly as a result of the fragmentation of data and applications in the local environment. Local data intermediaries, an institutional innovation in this field that has emerged since the early 1990s, are designed to address those barriers by assisting other local institutions in assembling community information and applying it productively. This section is largely based on the experience of the National Neighborhood Indicators Partnership (NNIP), introduced in chapter 1.

## Local Institutions and Community Information

All individuals are potential producers and users of community information. Through their lived experience, individuals learn new things about their community that guide their decisions. In this chapter, however, we are interested in roles played by local institutions that are responsible for the production and/or use of information. Figure 2.1 identifies the most important institutions in this regard. Below we review the types of institutions and the nature of, and trends in, their information-related functions.

### Local Government Agencies

Agencies of local general-purpose governments (i.e., counties and municipalities) are most important to this field, primarily because they produce most of the community information now available. Although community information can be derived from surveys, it is the automation of transactional data (particularly data produced by public agencies) that has been responsible for the revolution in the availability of community information.

**Figure 2.1.** Local Institutions and Their Roles in the Community Information Environment


Transactional data are produced when events, and descriptive information about those events, are recorded in administrative data files; for example, records produced on crime incidents, property sales, and births.[1] The automation of such data occurred because it promised important improvements in operational integrity and efficiency for local agencies, not to support community uses.[2] But once that automation occurred, the sheer volume of machine-readable community data in existence increased dramatically. Because the data were a by-product of the routine day-to-day work of administrative agencies, they could be reused for other purposes at low cost.

The advent and subsequent improvement of geographic information system (GIS) technology was the other critical innovation. The essay by Fleming at the end of this chapter explains the basics of geographic information systems. The essential feature is that all data are keyed to locations in a uniform manner,[3] allowing governments to layer diverse data.[4] Although data on properties and physical aspects of a community were the initial focus of most GIS applications, other types of address-based data (such as data on crimes, new business licenses, or the residences of public-assistance recipients) are now just as common.

The institutional implications of these changes within local governments have varied. The amount of automated information on social service clients has expanded markedly. Social service programs that offer case management may have electronic records on each of their clients, with information on what happened during and after each of their many interactions, as well as a host of other descriptive information about them. Agencies that maintain confidential information on individuals maintain close control of their own records, although efforts to integrate such data across programs have been under way recently in some places—see discussion of integrated data systems (IDS) in chapter 4.

Local government data that are a matter of public record (such as information on properties), however, have been more likely to be centralized institutionally. In her essay, Fleming notes

> Independent, stand-alone, single-department programs have given way to an enterprise approach whereby core datasets are developed, harmonized and made available for use by all service departments across the local government organization.

In earlier stages of automation, when records were stored on mainframe computers, department staff who needed to work with the data often faced long waits for limited hard copy reports from a central computer unit. An enormous boost in government productivity came when desktop computers were networked so staff could access the database directly to look up facts, perform analyses, and initiate appropriate updates. Many local governments have yet to adopt a full enterprise approach, but the trend is clearly toward more data sharing across departments. Chapter 3 discusses the more recent trend of local governments distributing data through centralized open data portals, which could spur a similar increase in governments' internal use of data.

Individual departments use the transactional records they produce in several ways. Maintaining official records with integrity is in itself a mission. The work entails not only updating records accurately as transactions occur (such as changing records on ownership when a property sale takes place), but making such information accessible to the public in an efficient way.

For individual departments, however, probably the most universal use of these data is basic program management: tracking to find out if and when assigned tasks have been accomplished. These data can also form the basis of more comprehensive management. In planning and implementing a department's activities more broadly, managers look at trends in the scale and composition of workloads and associated costs. Higher-level leaders in government use the same data in assessing the department's performance (see discussion of PerformanceStat in Fleming's essay following this chapter). Finally, the jurisdiction's elected officials or appointed managers, other top leaders, and central departments (such as the budget bureau and planning department) use data produced by their various departments, in combination with external data, for planning and management of the government's agenda overall, across departments.

## Other Individual Local Institutions

The list of other types of local institutions involved with community information is a long one. These local institutions both create data that can be turned into community information and make use of that data themselves in ways similar to those we have described for general-purpose local governments.

First, other public-sector institutions, most prominently school systems, keep detailed records. Virtually all school districts maintain regularly updated information on their individual students (e.g., test scores, absenteeism, graduation records), a practice highly valued for community indicators. Public hospitals and other public-sector health care providers are also in this category. They create a record when they provide service to someone, but those visits are sporadic and providers obtain minimal information about the people they serve.

Second are the numerous community groups and community-oriented nonprofit organizations that exist in all American cities. Importantly, these include community development corporations (CDCs) and other grass-roots neighborhood groups that work to improve community neighborhood well-being; as noted in chapter 1, these groups are primary clients for NNIP. This second group also includes nonprofit service providers as wide-ranging as child care centers, family financial counselors, community health clinics, charter schools, and homeless shelters. Groups that serve individuals are more likely to keep automated records on them now than they were a decade ago, but data quality is still likely to be less consistent than that of public agencies providing similar services.

CDCs, however, are more likely to have good data on the properties they own or are acquiring and rehabilitating.

For-profit companies commonly use community information for business development, including targeting marketing campaigns and selecting a site for a new store. The routine record keeping of some private firms can also yield data of value for public purposes. For example, data on the changing spatial pattern of a firm's product sales may indicate shifting demographic and income patterns that could serve as a basis for community group plans for facility expansion. Some firms now keep records on the residential addresses of their customers, which might provide fresh insights to transportation planners. There are many examples of this sort of application of private data for public purposes to date, and there could be more in the future.

Other institutions (and their purposes in data use) include philanthropies (plan and assess grant making); advocacy organizations (make cases and promote causes); some community-improvement nonprofits (plan and implement neighborhood initiatives); and research firms and university institutes (provide research and analysis for a variety of clients).

## Coalitions and Governance

So far we have talked about individual local institutions with roles in producing and using community information. But much of the collective decisionmaking and action at the local level necessarily involves a number of these institutions (arguably, an increasing share) working together. How does this work take place, and who is in a position to participate? The answer to these questions has a profound effect on the nature and extent of the demand for community information.

The answer begins by recognizing that the textbook picture of how collective decisionmaking works in a democracy at the local level is much oversimplified; that is, the notion that an election is held and the elected leaders then simply instruct the bureaucracy as needed to run the government until the next election. In fact, elections are only the starting point, and a healthy democracy requires that citizens participate actively in decisionmaking between elections.

How does this work in practice? Political scientists recognize that the process is one in which many groups with many objectives often compete for influence. There have been long-standing debates between those who see decisions being dominated by some type of elite and the pluralists who see access to decisionmaking as relatively open with more dispersed power (Judge, Stoker, and Wolman 1995).

Some who study the elites have seen the environment as highly competitive and emphasized ways in which business interests, development groups, and other coalitions of wealth have been able to exercise substantial control (Harding 1995). However, in the 1990s, a less hard-edged view emerged that is the most prominent paradigm in this field today: regime theory (Stone 1989, 2005). Regimes are governing coalitions of civic leaders that are collectively able to mobilize resources and implement action agendas of varying scopes.[5] The theory emphasizes "power to" (the ability to mobilize and implement agendas, recognizing that competition exists but relying more on collaboration than competition to achieve productive ends) rather than "power over" (the ability to unilaterally corral resources and control decisions in a more starkly competitive mode). Consistent with this theory, Briggs (2008) sees "democracy as problem solving." Stone (2005, 313) notes

> A governing coalition consists typically of members based in a locality's major institutions. It is a mistake, however, to think of urban regimes as composed of a fixed body of actors taking on an ever-changing agenda. Instead, the question is about who needs to be mobilized to take on a given problem effectively . . . the issue addressed determines whose participation is needed.

But he also recognizes that regimes organize around larger purposes and have continuity (Stone 2005, 318)

> The framework of durable purpose means that governing arrangements do not have to be reinvented issue by issue . . . a big-purpose agenda can minimize transaction costs by providing established and familiar ways of getting things done. . . . It does so on a basis that gives structure to interactions over time.

As to the substantive scope of the work, regime theory emphasizes governance rather than government. The former encompasses the very broad set of institutions, approaches, and norms involved in managing collective life; the latter is the more narrowly prescribed institutional and procedural mix officially authorized for conducting public business.[6]

The idea of the governing coalition clearly implies substantial involvement by a wide group of players in public policy decisions, including civil society. Key members of the coalition are consulted and involved in negotiations. Leaders might include the chair of the community foundation, a local university president, the head of the United Way, nonprofit leaders, and so forth, along with government officials.

Community information can be valuable in almost all phases of civic deliberations, as well as in informing decisions about smaller issues. How data can be used in these processes is discussed more in chapters 4, 5, and 6, but the implication is that the range of local users is broad indeed. They may include ongoing regimes that are furthering a number of broad longer-term issues (such as school reform), but there are also smaller coalitions that focus on goals like improving the quality of life in particular districts or neighborhoods.[7]

## State and Federal Institutions

State and federal institutions are not direct players in the local decision-making milieu, but they certainly influence community information at the local level. First, they develop and publish data files with information on small areas. A great benefit of these files is that they employ consistent variable definitions and protocols state- or nationwide that enable comparisons across jurisdictions. These data sources are reviewed in chapter 3.[8] Federal examples include Decennial Census and Home Mortgage Disclosure Act data.

At the state level, data are collected on programs that are state supervised or administered, such as child welfare and public assistance programs. In one prominent example, vital statistics are often maintained by state health departments. Although the records in these programs originate at the local point of contact, they are typically managed through statewide information systems. Local officials may be able to access summary reports from these data repositories, but they do not have the direct control over the records as they do in locally administered data systems.

Second, where they have jurisdiction to do so, higher-level governments regulate the way some kinds of data can be used locally. This is particularly true of confidential information about individuals and families. There are always restrictions on the use of these data and their release to outside parties. Agencies have long devised their own restrictions related to these people-referenced records, but 15 federal agencies follow the specifications of the 1991 Federal Policy for the Protection of Human Subjects, also known as the Common Rule. The policy protects human subjects involved in any research conducted, supported, or regulated by the agencies (Petrila 2011). Two other important pieces of federal legislation clarified uses of particular types of data: the Health Insurance Portability and Accountability Act (HIPAA) and the Family

Educational Rights and Privacy Act (FERPA) (Farley and Polin 2012; Petrila 2011). Although the rules constrain the use of data of this type, they certainly have not prevented it.

Third, state and federal actors often provide strong financial or other incentives to local agencies to encourage them to maintain administrative records and to use their data actively to manage programs effectively. Such encouragement has expanded over the past decade. One important recent example is the American Recovery and Reinvestment Act provision, which specified that states must provide assurance they will establish a unified longitudinal (P-20) data system on education to receive state fiscal stabilization funds. These systems are created by recurrently assembling student-level data from local school systems, and provisions call for developing the capacity to link preschool, K–12, and postsecondary education, as well as workforce data (with appropriate confidentiality protections); see, for example, Data Quality Campaign (2012).

Finally, state and federal governments enact new programs that encourage local actors to take on additional activities or modify their processes. These may, in and of themselves, put pressure on local agencies to change how they develop and use community information, simply because enhanced information capacity is implicitly needed to be able to implement the new approach effectively (or at all). However, specific requirements pertaining to information may also appear in program legislation or regulations.

The Obama administration has been particularly active in this regard. An important example is its Promise Neighborhoods initiative, which offers a "cradle to career" approach to child and youth development (Promise Neighborhoods Institute 2011; Biglan et al. 2011). This competitive grant program seeks a coherent set of interventions that begins with early childhood development and stretches into linkages to fruitful careers, all focused on the children within a defined neighborhood boundary. As discussed in chapter 5, the capacity to use information in planning, tracking, and performance management is prominent in the criteria for making awards and in subsequent federal oversight of implementation (Comey et al. 2013).[9]

## Local Data Intermediaries: An Institutional Innovation

The preceding section showed that a vast amount of community data is now being produced in American localities with many types of would-be users, ranging from individual small agencies to large mission-driven coalitions. Interest has grown as many local stakeholders recognize that they have no solid basis for targeting resources or evaluating programs created to improve conditions if they lack hard facts on the neighborhood-level patterns and trends.

Program operators can often learn much just by analyzing their own data (the administrative data they produce) more effectively. But almost all of them need access to a variety of neighborhood-level trends to understand the dynamics of the neighborhood context in which their programs operate. Police departments, for example, now clearly recognize that they cannot deploy crime prevention resources effectively by simply analyzing data on crimes. They need to look at a number of other indicators as well before they can understand a neighborhood and how it may be changing.

The problem is that the trend toward government enterprise data systems has still not reached many communities, and in many cities, the data of interest are currently stored in individual agency database silos. Community groups may recognize the need for cross-topic neighborhood-level data, but it would obviously be extremely wasteful for all such groups to go from agency to agency to try to collect the woefully inadequate data typically being released to the public. Their time would be much better spent on activities to further their own missions.

A more efficient approach is to assign the data assembly task to a local data intermediary—one single institution or a formal collaboration of institutions—that will assemble and organize the data across local government agencies and build a system to serve as a one-stop shop for community information. The goal is an entity that will provide accurate and useful data on multiple topics to all groups that need it and that will commit to continuing the data provision over the long term. These are the kinds of entities and collaborations that have become a part of the National Neighborhood Indicators Partnership (NNIP), introduced in chapter 1.

Where such intermediary capacities have been developed, interviews suggest that they are regarded as valuable by community groups, which previously had hardly any access to relevant data. They are also valued by public agencies and nonprofits that need a richer understanding of neighborhood conditions and trends to plan their own work effectively, but lacked the internal capacity to perform their own cross-agency data assembly and analyses. Because the data are regularly updated, users do not have to start all over again with a long period of data collection when they want to update a previous study of an issue. Building an adequate system necessarily entails some cost, but it is almost sure to represent a net savings compared with the resources so many local groups now spend trying to collect data with such unsatisfying results. Leaders, including local philanthropies, see the cost-effectiveness of investment in local professionals who understand their environment and will be there over time to interpret and advise people who live in their cities. The remainder of this section explains more about institutions that have taken on this local data intermediary role and discusses their three basic functions in more detail.

## Types of Institutions Serving as Local Data Intermediaries

As the logic of having a local data intermediary becomes understood in a city, NNIP experience suggests that one or a small number of local entities (e.g., university institutes or nonprofits with data expertise) commit to developing the capacity to perform this work. Civic leaders (i.e., governing coalitions) mobilize the core funding for local data intermediaries and agree on the local institutions to serve in that role. Two features of these selections are noteworthy.

First, most NNIP partners are outside of government. One might think a municipal agency (probably the city planning department) would be selected for this work, but that has not happened so far in NNIP experience. Two reasons explain the benefits of having a nongovernmental organization function as the data intermediary. Some city charters do not permit their agencies to perform the full range of data intermediary functions. In addition, government agencies are seen as more likely to be responsive to current elected officials than to broader and longer-term communitywide interests. Nongovernmental organizations are less likely to be "owned" by the current mayor or any political faction, either in perception or reality. This neutral reputation of nongovernmental groups enables them to more easily obtain data from, and develop trusted working relationships with, a wide variety of data providers and users in and outside of government. Such relationships can last over the long term.

The second feature of interest is the variety of institutional types that has been selected. NNIP experience suggests that basic institutional type is less important in the selection than other features. The local organization must have strong leadership and technical skills. It should also be viewed as credible by its diverse audiences, a reputation that is earned by consistently providing unbiased data and analysis. The NNIP partners illustrate the range of institutions that demonstrates these characteristics and can play the role of local data intermediary (see table 2.1 and figure 2.2).[10]

Among NNIP partners, 29 are individual institutions. Of these,

- nine are community-oriented university departments or research centers;
- six are freestanding nonprofits that perform data intermediary work exclusively;
- eight are subunits of freestanding nonprofits that perform the data intermediary work along with broader community improvement or direct service missions;
- three are government agencies (a library, a public health department, and a regional planning agency); and
- three are local funders (e.g., a community foundation).

The remaining seven partners are formal collaborations of multiple local institutions in the above categories, with five of them a partnership of two organizations. Several collaborations involve partnerships between an entity based in the central city (nonprofit or university center) and the planning agency for the metropolis as a whole. This collaborative arrangement facilitates the presentation and use of data at differing scales.

Some of the strongest NNIP partners are freestanding organizations whose work focuses only on the data intermediary functions described earlier in this section. However, other NNIP partners are a program within a large local institution (such as a multipurpose nonprofit, a university, or a community foundation) rather than a freestanding entity. The main benefit of this model is the prospect of sustainability over the long term. Program sustainability derives in large part from the stability of the underlying institution itself, but it also derives from the possibility of financial benefits (some overhead and in-kind support) the larger institution may be able to provide.

Local institutions that form coalitions to join NNIP usually come together because when combined, their differing capacities allow them to perform all the data intermediary roles. For example, one institution may specialize on the information system side, while another specializes

**Table 2.1.** National Neighborhood Indicators Partner Organizations, January 2014

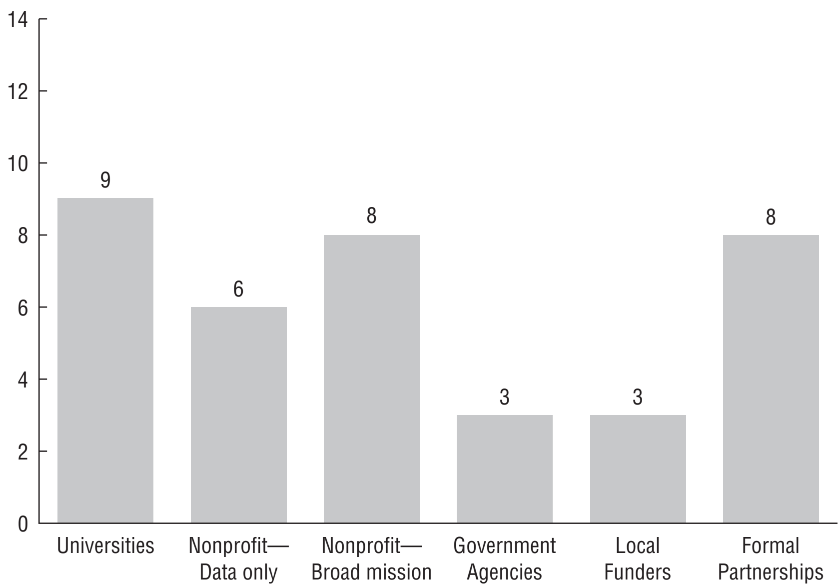
| City/Metro | Organization |
|---|---|
| Atlanta, Georgia | Neighborhood Nexus |
| | Community Foundation for Greater Atlanta |
| | Atlanta Regional Commission |
| | Emory University Office of University-Community Partnerships |
| Austin, Texas | Children's Optimal Health |
| Baltimore, Maryland | Baltimore Neighborhood Indicators Alliance, University of Baltimore |
| Boston, Massachusetts | The Boston Foundation |
| | Metropolitan Area Planning Council |
| Camden, New Jersey | CamConnect |
| Chattanooga, Tennessee | Ochs Center for Metropolitan Studies |
| Chicago, Illinois | Chicago Metropolitan Agency for Planning |
| Cleveland, Ohio | Center for Urban Poverty and Social Change, Case Western Reserve University |
| Columbus, Ohio | Community Research Partners |
| Dallas, Texas | Institute for Urban Policy Research, University of Texas at Dallas |
| Denver, Colorado | The Piton Foundation |
| Detroit, Michigan | Data Driven Detroit |
| Des Moines, Iowa | United Way of Central Iowa; Child and Family Policy Center |
| Grand Rapids, Michigan | Community Research Institute, Grand Valley State University |
| Hartford, Connecticut | HartfordInfo |
| Indianapolis, Indiana | Polis Center, Indiana University–Purdue University Indianapolis |
| | United Way of Central Indiana |
| Kansas City, Missouri | Center for Economic Information, University of Kansas City–Missouri |
| | Mid-America Regional Council |
| Louisville, Kentucky | Network Center for Community Change |
| Memphis, Tennessee | Center for Community Building and Neighborhood Action, University of Memphis |
| Miami, Florida | The Children's Trust |
| Milwaukee, Wisconsin | The Nonprofit Center of Milwaukee |
| Minneapolis–St. Paul, Minnesota | Center for Urban and Regional Affairs, University of Minnesota |

*(continued)*

**Table 2.1.** *(Continued)*

| City/Metro | Organization |
|---|---|
| Nashville, Tennessee | Neighborhoods Resource Center |
| New Haven, Connecticut | DataHaven |
| New Orleans, Louisiana | Greater New Orleans Community Data Center |
| New York, New York | Furman Center for Real Estate and Urban Policy, New York University |
| Oakland, California | Urban Strategies Council |
| Philadelphia, Pennsylvania | Metropolitan Philadelphia Indicators Project at Temple University |
| | The Reinvestment Fund |
| Pinellas County, Florida | Juvenile Welfare Board |
| Pittsburgh, Pennsylvania | University Center for Social and Urban Research, University of Pittsburgh |
| Portland, Oregon | Institute of Portland Metropolitan Studies, Portland State University |
| Providence, Rhode Island | The Providence Plan |
| Sacramento, California | Community Link Capital Region |
| St. Louis, Missouri | Rise |
| San Antonio, Texas | CI:Now |
| Seattle, Washington | Public Health—Seattle and King County |
| Washington, D.C. | NeighborhoodInfo DC |
| | Urban Institute and LISC |

**Figure 2.2.** Number of NNIP Partner Sites by Institutional Type, January 2013



in hands-on data work with communities or, as already noted, one handles central city data while another works at the regional scale.

Many workable variations are possible. Large cities, in particular, may require creative thinking about coalition building because many institutions with various data capacities may already be in place. A recent review for Chicago, for example, concluded that the region already had many strong data-oriented institutions that could fulfill a large part of the overall intermediary role and that no new central intermediary entity was required. Rather, the authors suggested that the existing institutions join in a formal network that would keep all members informed of relevant new developments and would work in a coordinated way to address local information needs (Pettit and Kingsley 2013).

Once a local data intermediary begins operating in a collaborative way, its work is likely to strengthen interests in data and data capacities of many other local institutions. In their essay accompanying this chapter, Wascalus and Matson tell how a particularly rich set of ongoing institutional relationships has developed around community information in metropolitan Minneapolis–Saint Paul. The NNIP data intermediary is an important participant, but it works with several other local research groups as well as a sizable number of community and regional players, with the mix adjusting as different issues are on the table.

### *Commitment to Distressed Neighborhoods*

Regardless of the local institutional arrangement, the NNIP partner organizations share a commitment to the value that motivated the creation of NNIP: to use data to empower and improve the conditions in distressed neighborhoods. All NNIP partners employ two fundamental strategies to fulfill this obligation. First, they provide data and technical assistance to advocacy, service, and resident groups located in low-income neighborhoods, as well as citywide groups focusing on improving opportunities in those places. The access to information can help level the playing field by giving people who were previously shut out of public policy and planning decisions the capacity to knowledgeably participate in the conversation. This is a central example of partners' contribution to strengthening civic life and governance (described below). Second, the partners perform and disseminate analysis on topics related to low-income households and neighborhoods. The focus can be overt, in studies on topics such as child poverty, or less direct, through work on broader topics like economic development or education policy.

The partners' adoption of additional methods to strengthen distressed neighborhoods depends on their core mission. Some organizations, such as the Greater New Orleans Community Data Center, commit to supporting inclusive communities, but function as explicitly neutral providers of data and analysis. Other advocacy and organizing groups then use the data to promote equitable policies and mobilize disadvantaged communities. In contrast, some partner organizations are vocal advocates for social justice. For example, the Urban Strategies Council website states that "all policies, practices, and programs must be considered through an 'equity lens' where eliminating disparities is a primary concern" (Urban Strategies Council n.d.). Their "equity model" includes data collection to support equity, as well as the responsibility of defining and communicating equity and ensuring an initiative's leadership is representative of community. Wherever NNIP partners fall on the spectrum of activism, they all must be sensitive to their relationships with government agencies as providers and users of the data.

### *Funding and Sustainability*

NNIP partners are varied in scale. A 2009 survey of 24 NNIP partner organizations showed an average annual budget for data and information services of $335,000, ranging from $75,000 to $1.1 million.[11] The total was below $150,000 for 38 percent of the partners and above $500,000 for 29 percent. They had on average 4.1 full-time equivalent employees.[12]

All partners received some general-support funding. Local philanthropies and the broader institutions of which NNIP organizations are a part provided most of the funding of that type (e.g., support provided by a university to its own community research institute). However, all the partners are also funded in part by fees received for conducting studies or performing other data-related services for various clients. Such fees accounted for 58 percent of total revenues on average, but there was considerable variation in that amount. Review of the information on scale and funding composition revealed no systematic differences between university-based intermediaries and those housed in other types of institutions. There was considerable variation within both groups.

There is a sizable fixed-cost component in expenses for basic data intermediary functions (assembling and cleaning data from a number of sources, entering them into an orderly system, and releasing them to the public over the web in various forms). These fixed costs do not vary much with the size of the city.

NNIP experience suggests that operating a local data intermediary is financially sustainable. The growth of NNIP partners from 6 to 37 locales shows that fundraising for local data intermediary functions is possible in a variety of governmental and philanthropic contexts. As shown above, success in these activities has also been achieved under a variety of institutional forms. The NNIP partners still face the same funding challenges as any nongovernmental organization providing a community service. Although their financial strength varies, all share difficulties in raising money for basic data system development and for staff time to be responsive to requests from community groups who cannot afford to pay. They must stay attuned to shifts in the funding environment and continually demonstrate their value to the foundations and government agencies that support them.

Nonetheless, NNIP partners have been able to navigate their local environments and sustain operations for many years almost entirely with local funding. Nine have been in operation for 15 years or more, and none relies on sole support from a national philanthropy or the federal government. Since the NNIP network was founded in 1995, several partner organizations have closed, but in almost all cases, the NNIP functions were successfully transferred to and sustained by another local institution. The longevity of the NNIP partner organizations validates the network's efforts to present the model to other cities and metropolitan areas as a viable approach for expanding the availability and use of community information.

## The Functions of Local Data Intermediaries

Local data intermediaries perform three basic functions: (1) assembling, transforming, and disseminating data; (2) applying the data to achieve impact; and (3) using data to strengthen civic life and governance.[13] These functions describe how data intermediaries work in most cities and how communities are altered by the presence of a data intermediary. Although the main functions are the same, many aspects of the way they are performed have changed, particularly in response to the remarkable technical advances and increases in data availability that have taken place over the past 15 years (these changes are addressed in chapter 3).

### *Assembling, Transforming, and Disseminating Data*

*Data assembly.* This most basic function for local data intermediaries entails obtaining data through open-data portals or data-sharing agreements. The intermediaries commit to regularly updating the data over the long term; the commitment of civic leaders to support this activity needs to be long term as well. With this model, local stakeholders only have to go to one source to access neighborhood data on a variety of topics.

Assembling data across topics leads to greater insights. For example, foreclosure and home sales data are critical to demonstrating the housing market impacts of the foreclosure crisis across neighborhoods. However, by linking those data to school enrollment data and crime data, as some NNIP partners have done, it becomes possible to identify the human consequences of the crisis as well. These data then provide the basis for consideration of a wider range of policy responses, such as changing policies on school assignment to reduce student school mobility due to involuntary moves (Pettit and Comey, 2012).

*Data transformation.* Even for experts, working with raw administrative data to create useful measures is challenging and can be very costly, especially when it becomes necessary to combine data from the files of different agencies. Intermediaries play an essential role in transforming data to make them easier for nontechnical users to understand. This work includes cleaning the data and creating new indicators, new forms of display, and metadata (e.g., definitions and documentation on processing).

Because data intermediaries regularly update datasets they have worked to obtain, they build up substantial knowledge over time about the reliability of the data and the purposes for which they are best used. This knowledge is used to improve the transformation process. The example described above of the analysis of linked school enrollment and foreclosure rates requires substantial transformation of data from each source.

Government datasets released through open-data portals support many important policy and service innovations today, and they will be the basis for more in the future. But most of these data will be in a raw form that will be difficult for those without a high level of expertise to use directly. These datasets still require transformation to maximize their usefulness for influencing policy and improving communities.

In addition, when contributing agencies update their administrative datasets internally, they often replace the old version with the new so that data on past transactions are lost. Intermediaries sometimes take on the role of archiving the historical data, creating longitudinal datasets that enable the examination of trends over time.

*Disseminating data.* For the most part, dissemination means releasing the data directly to the public over one or more websites in different forms to suit a range of needs and technical expertise. Dissemination by intermediaries ranges from publishing static displays, such as maps and statistical profiles with charts and tables for individual neighborhoods, to providing structured data files that users can download. Some displays are interactive, and users can specify aspects of the form and content of the maps or charts they want. On some websites, for example, a user can click on a particular land parcel on a map, and the system then brings up various displays about that parcel, including a photograph of the building and tables with descriptive characteristics. Data intermediaries also disseminate data in more traditional forms, such as hard copy reports and fact sheets. As described below, a large part of the dissemination role that data intermediaries play in their community is in working directly with local stakeholders, engaging them in using the data to influence policy and achieve impact.

## *Applying the Data to Achieve Impact*

Although the acquisition, transformation, and dissemination of data are essential, NNIP partners consider their most important function to be applying the data to address local policy problems. Data and technical assistance from data intermediaries can motivate changes to public policies at the neighborhood, local, and state levels. The intermediaries endeavor to increase the ability of governments and community organizations to identify emerging issues; to target resources and investments efficiently; and to empower groups in distressed neighborhoods to mount improvement initiatives.

Clients include government agencies, city councils, community foundations, nonprofit service providers, neighborhood associations, and community development corporations. The intermediaries often prepare analyses and reports on particular topics themselves. The experiences can be even more beneficial, however, when the intermediary works interactively with individual clients and helps them to understand

and use the data. In these engagements, they prompt learning along the way so that at the end, the clients legitimately feel that they "own" the final products. This emphasis, particularly when working with low-income neighborhoods, gives rise to NNIP's use of the term *democratizing information.*[14] These interactions can be extensive, but many local data intermediaries also run help desks that local groups can call for limited one-on-one help with data tasks.

Many times, simply the release of data, such as issuing a news release or brief report showing some surprising new trend, can have a large impact on the local policy environment. However, most of the work of local data intermediaries is more complex. Examples include using the data to inform the design of a neighborhood improvement initiative; working with funders to conduct community needs assessments; helping an individual agency or nonprofit use data in program planning and performance management; or conducting independent program evaluations. Additional examples have been documented in earlier NNIP publications (Cowan 2007; Kingsley et al. 1997), and more are presented in chapters 5 and 6.

### *Using Data More Broadly to Strengthen Civic Life and Governance*

In many cases NNIP partners have had important impacts (1) by providing general advice, technical assistance, and training to government agency staff and community practitioners to help them build internal data capacity and (2) by working directly to improve the quality of data systems maintained by public agencies. The direct engagement work of NNIP partners with local organizations around data and analysis helps the staff or residents become more savvy users of data and information. In addition, a number of NNIP partners provide training on specific data sources, the use of data, and the local intermediary data portals. For example, Rise in St. Louis, Missouri, holds trainings to teach local community development corporations how to use and interpret data from the American Community Survey. Several partners also offer training on using their online data query and visualization systems. For example, the Polis Center at Indiana University–Purdue University in Indianapolis, Indiana, offers training courses centered on the use of their SAVI system to address local issues.[15] An example with broader objectives is the Certificate in Program Design and Development offered by the Institute for Urban Policy Research at the University of Texas at Dallas. In a series of five classes, nonprofit organizations learn how to integrate their theory and change, logic model, and program design elements to increase effectiveness and document outcomes.

Beyond strengthening the data capacities of local institutions, local data intermediaries have the potential to play an important role in changing the culture, through developing a community of practice among local stakeholder organizations to promote the effective use of data in decisionmaking. This role may include the intermediary convening regular meetings at which all participants can share innovative applications, identify gaps in local practice and ways to address them, and build a constituency for productive data efforts.

In line with these objectives, there is the potential for local intermediaries to devote more time to partnering with outside experts. For example, they might provide data to and collaborate with representatives of the Center for Community Progress, invited to their city to advocate for property reclamation and offer related technical assistance. Alternatively, they might partner with technology firms that help local stakeholders develop new software applications or assemble and use data in other creative ways. Data intermediaries can encourage local leaders to bring in such groups and collaborate in the process to enhance payoff and productive "leave-behinds." The growth of the open data movement, discussed in chapter 3, offers one such opportunity. Pettit et al. (2014) documents the roles NNIP partners currently play in this area and reflects on ways local data intermediaries can work with open data advocates in furthering the goals of broad access to information.

## The NNIP Network

NNIP has operated since its formation in 1996, with considerable growth in membership. The focus of this book is on work at the local level, but this brief introduction to the network illustrates one source of national support to expand and improve the local work.

An executive committee, elected by and from the local partners, is central to planning and overseeing the work of the NNIP network overall. The Urban Institute serves as NNIP's secretariat and works closely with the executive committee in planning and implementing activities. The five parts of NNIP's overall agenda are summarized below and explained in more depth in Kingsley, Pettit, and Hendey (2013).

1. *Informing local policy initiatives.* NNIP's cross-site action initiatives are applications of data designed to help address real local issues, but they are structured in a comparable manner in multiple NNIP cities so as to provide lessons for national, as well as local, policy and practice. NNIP coordinates the local work and documents best practices to guide other cities interested in working on the topic. Example topics have included neighborhood public health, reintegrating released prisoners into society, decision-support tools for community development, and early childhood development and school readiness.

2. *Developing tools and guides.* This entails preparing and disseminating guidebooks, tools, and presentations that advance the state of the art in this field. Topics range from descriptions of promising practices developed in cross-site initiatives to technical guidebooks documenting specific datasets.

3. *Strengthening local capacity: developing capacity in new communities.* New cities learn about NNIP through the website, national presentations, and the work of local partners. Urban Institute staff offer limited technical assistance to cities interested in starting a local data intermediary if needed, and assist the groups in formally applying for membership.

4. *Strengthening local capacity: services to an expanding network.* The network's most important mechanism for achieving its objectives is peer-to-peer learning among NNIP partner organizations. This is implemented through two face-to-face meetings of the full partnership each year and informal activities between meetings. The network's website is also used to disseminate information to broader audiences interested in this work.

5. *Leadership in building the field.* NNIP works to catalyze a broader effort to promote local use of community information in decisionmaking by partnering with other national organizations whose missions revolve around the use of indicators and development of local capacity. Groups such as the Federal Reserve Board of Governors, the National League of Cities (NLC), and the International City/County Managers Association (ICMA), have convened stakeholders and provided trainings related to local data use in partnership with NNIP.

## Conclusions

Three conclusions stand out as most important in considering how best to enhance the role of community information in making decisions about collective issues in local governance.

First, the number of institutions that need to be involved is likely to be large, and their interactions will inevitably be complicated. In this field, trying to work only with local government or any other single institution, even on a single relevant issue, is hardly ever sufficient. Fortunately, most urban areas already have networks of community groups and coalitions of civic leaders that offer good places to start more productive engagements.

Second, the automation of administrative records by these institutions for their own individual purposes has vastly expanded the amount of community information that is now potentially available. Most of these institutions are both producers and users of this information, and it is important to remember that they play both roles.

Third, in most urban areas, the data remain in individual agency silos; that is, they are not shared across agencies. However, in a significant number of places, local entities—local data intermediaries—have been formed that have learned how to break down those silos effectively. They have been able to convince local agencies to share their data, demonstrating how doing so can expand the range of data available for agencies' own purposes as well as support broader collective applications.

What kinds of applications have been implemented so far to illustrate the value of this approach? Various examples are described in chapters 5 and 6. However, for the reader to understand them, we must first present a more complete picture of recent advances in data availability and technical capabilities (chapter 3) and of the basic types of applications that are relevant in community work and how data are used within them (chapter 4).

### NOTES

1. The types of local transactional data now available are reviewed in chapter 3. See also the catalog of administrative data sources for neighborhood indicators provided in Coulton (2008).

2. Today, most local government records are automated. The current central handbook of the International City/County Managers Association on record keeping is entitled *Electronic Records Management;* see Stephens (2005).

3. Location is almost always defined first by geospatial (latitude–longitude) coordinates, but it may also be linked to polygons such as parcel layers containing the boundaries of land as defined by the local government.

4. See also Fleming (2005) and Pettit et al. (2011).

5. As Stoker (1995) states, "In a complex society, the crucial act of power is the capacity to provide leadership and a mode of operation that enables significant tasks to be done. This is the power of social production."

6. Barnes and Foster (2012) propose a set of dimensions and factors of regional governance that focuses on capacity and purpose.

7. See Martin and Morehead (2013) for a case study on how the Greater Portland Pulse community indicators projects contribute to various aspects of regional governance.

8. See also Coulton (2008).

9. The Promise Neighborhoods initiative is only one of several new federal programs that focus on local development and provide similar encouragement for advanced development and use of information. Others, which are discussed in chapters 5 and 6, include Sustainable Communities (US Department of Housing and Urban Development 2012) and its umbrella, the Neighborhood Revitalization Initiative (White House 2011).

10. For additional information about NNIP partners and the network, see www.neighborhoodindicators.org.

11. Partners estimated the dollar amount that was spent on data intermediary functions, not the entire organizational budgets.

12. For additional information, see http://www.neighborhoodindicators.org/partners/about-our-partners.

13. The paragraphs describing the functions of local data intermediaries are adapted from a strategic review conducted in 2012 (Kingsley, Pettit, and Hendey 2013).

14. Sawicki and Craig (1996) applied this term as a central NNIP principle and provide a longer exploration of this topic.

15. For additional information, see http://www.savi.org/savi/training.aspx.

## REFERENCES

Barnes, William R., and Kathryn A. Foster. 2012. "Toward a More Useful Way of Understanding Regional Governance." Building Resilient Regions working paper. Presented at the conference of European Urban Research Association, Vienna.

Biglan, Anthony, Christine Cody, William Aldridge II, Alexis Dubroski, and Jean Kjellstrand. 2011. "The Promise Neighborhoods Initiative: Improving Developmental Outcomes through Comprehensive Interventions." *Journal of the Institute for Comprehensive Community Development* 2 (2): 25–33.

Briggs, Xavier De Souza. 2008. *Democracy as Problem Solving: Civic Capacity in Communities Across the Globe.* Cambridge, MA: MIT Press.

Comey, Jennifer, Peter A. Tatian, Lesley Freiman, Mary K. Winkler, Christopher Hayes, Kaitlin Franks, and Reed Jordan. 2013. *Measuring Performance: A Guidance Docu-*

*ment for Promise Neighborhoods on Collecting Data and Reporting Results.* Washington, DC: Urban Institute.

Coulton, Claudia J. 2008. *Catalog of Administrative Data Sources for Neighborhood Indicators.* Washington, DC: Urban Institute. January.

Cowan, Jake. 2007. *Stories: Using Information in Community Building and Local Policy.* Washington DC: Urban Institute.

Data Quality Campaign. 2012. *Data for Action 2012: Focus on People to Change Data Culture.* Washington, DC: Data Quality Campaign.

Farley, Chelsea, and Meredith Polin. 2012. *Collective Ideas to Collective Impact: A Guide to Data Collaboration in Communities.* Palo Alto, CA: nFocus Solutions.

Fleming, Cory. 2005. *The GIS Guide for Local Government Officials.* Redlands, CA: Esri Press.

Harding, Alan. 1995. "Elite Theory and Growth Machines." In *Theories of Urban Politics,* edited by David Judge, Gerry Stoker, and Harold Wolman. Thousand Oaks, CA: Sage Publications.

Judge, David, Gerry Stoker, and Harold Wolman. 1995. "Urban Politics and Theory: An Introduction." In *Theories of Urban Politics,* edited by David Judge, Gerry Stoker, and Harold Wolman. Thousand Oaks, CA: Sage Publications.

Kingsley, G. Thomas, Claudia J. Coulton, Michael Barndt, David S. Sawicki and Peter Tatian. 1997. *Mapping Your Community: Using Geographic Information to Strengthen Community Initiatives.* Washington, DC: U.S. Department of Housing and Urban Development, October.

Kingsley, G. Thomas, Kathryn L.S. Pettit, and Leah Hendey. 2013. *Strengthening Local Capacity for Data-Driven Decisionmaking.* Washington, DC: Urban Institute.

Land, Kenneth C. 1975. "Social Indicator Models: An Overview." In *Social Indicator Models,* edited by Kenneth C. Land and Seymour Spilerman. New York: Russell Sage Foundation.

Martin, Sheila A., and Elizabeth Morehead. 2013. "Regional Indicators As Civic Governance: Using Measurement to Identify and Act Upon Community Priorities." *National Civic Review* Spring: 33–42.

Petrila, John. 2011. *Legal Issues in the Use of Electronic Data Systems for Social Science Research.* Philadelphia, PA: Actionable Intelligence for Social Policy, University of Pennsylvania. http://www.sp2.upenn.edu/aisp_test/wp-content/uploads/2012/12/0033_12_SP2_Legal_Issues_Data_Systems_000.pdf.

Pettit, Kathryn L. S., and G. Thomas Kingsley. 2013. *An Assessment of the Chicago Metropolitan Area Community Information Infrastructure.* Washington, DC: Urban Institute.

Pettit, Kathryn L. S., and Jennifer Comey. 2012. *The Foreclosure Crisis and Children: A Three-City Study.* Washington, D.C.: The Urban Institute.

Pettit, Kathryn L. S., Leah Hendey, Brianna Losoya, and G. Thomas Kingsley. 2014. *Putting Open Data to Work for Communities.* Washington, DC: Urban Institute.

Pettit, Kathryn L. S., and G. Thomas Kingsley. 2011. "Framework: The New Potential for Data in Managing Neighborhood Change." In *Putting Data to Work: Data-Driven Approaches to Strengthening Neighborhoods.* Washington, DC: Federal Reserve Board of Governors. December.

Pettit, Kathryn L. S., Barika Williams, Peter Tatian, and G. Thomas Kingsley. 2011. *Using Property Databases for Community Action.* Washington, DC: Urban Institute.

Promise Neighborhoods Institute. 2011. *What is a Promise Neighborhood?* Oakland, CA: Promise Neighborhoods Institute. http://www.promiseneighborhoodsinstitute.org/What-is-a-Promise-Neighborhood.

Sawicki, David S., and William J. Craig. 1996. "The Democratization of Data: Bridging the Gap for Community Groups." *Journal of the American Planning Association* 62 (4).

Stephens, David O. 2005. "Electronic Records Enter the Mainstream." In *Electronic Records Management,* edited by Robert Freeman, Mark R.W. Lemahieu, Jason Sanders, David O. Stephens, and Julian L. Mims III. Washington, DC: International City/County Managers Association.

Stoker, Gerry. 1995. "Regime Theory and Urban Politics." In *Theories of Urban Politics,* edited by David Judge, Gerry Stoker, and Harold Wolman. Thousand Oaks, CA: Sage Publications.

Stone, Clarence. 1989. *Regime Politics: Governing Atlanta.* Lawrence, KS: University Press of Kansas.

———. 2005. "Looking Back to Look Forward: Reflections on Urban Regime Analysis." *Urban Affairs Review* 40 (3): 309–41.

Urban Strategies Council. n.d. The Equity Framework. Accessed February 16, 2014. http://www.urbanstrategies.org/equity/#.UwEoIWCYaM8.

US Department of Housing and Urban Development. 2012. *Office of Sustainable Housing and Communities: Helping Communities Realize a More Prosperous Future.* Washington, DC: US Department of Housing and Urban Development. November.

White House. 2011. *Building Neighborhoods of Opportunity: White House Neighborhood Revitalization Initiative Report.* Washington, DC: The White House. July.

ESSAY

# Technology, Data, and Institutional Change in Local Government

Cory Fleming

Local governments exist to serve the needs of their communities. One of those needs involves the collection and maintenance of data related to the community. Many local governments have recognized the value of these data and have begun to mine the data for critical business intelligence. As a result, local governments have begun to change the way they do business. This essay considers a few key technologies that have been widely adopted by local governments and highlights how the resulting data generated by these technologies are driving institutional change in local government.

## Background

Historically, local governments have maintained many of the official records of their communities. This duty, particularly the maintenance of property records, has been critical for the development and evolution of communities and their economies. The existence of property records at local government offices gave landowners clearly defined legal rights and encouraged them to invest and make improvements to their land, which in turn fostered growth and development in the community. The records maintained by local governments have long been recognized as authoritative sources of data and information for many purposes.

Technology has greatly changed how local governments collect and maintain data. In the span of less than three decades, offices have moved from working with mainframe computers and electric typewriters to laptops, tablets, and smartphones. As old paper records have been scanned and converted to digital form, local governments have been able to manipulate and analyze data in a variety of new ways. More powerful and affordable computer systems have made it possible to work with increasingly large amounts of data. The need to hire staff to collect and process data is being replaced by the need to hire data analysts to probe available data and answer policy questions.[1]

A renaissance is taking place. Web and mobile applications, sometimes referred to as Web 2.0, permit ever larger amounts of data to be collected, shared, and analyzed, pushing community leaders to rethink how business should be done. As a result, communities are changing in the way they develop and even what it means to be a community. For community leaders—and local governments more specifically—the ability to connect with citizens increased remarkably with the introduction of websites and e-mail. Communications and interactions with the public have become more immediate and interactive as a result, making it easier for leaders to better understand what issues are of greatest importance to citizens.

Currently citizens can access nonconfidential municipal documents almost any time they want via government websites. In the not too distant future, smartphone users will be able to hear and see the public discussions relating to their property or proposal, as well as view all related regulations, plans, and documents. The need to attend a public hearing in person will be replaced by virtual meetings using video-conferencing technology. Data centers with in-house servers are being reconfigured with virtual servers (more popularly known as the cloud). The future will bring greater integration of all data, including video and voice. By virtue of their historic role as keeper of a multitude of administrative records, local governments need to be at the forefront of this effort. Indeed, the data and information systems used for maintaining and updating these records are considered essential infrastructure for the community.

Local governments are also demonstrating that the analysis of data can be used to realize significant return on investment, cost savings, and better decisions. With the ability to better gauge the potential results of policy decisions, local governments are embracing new innovations for creating neighborhood and community change. They are also moving away from their role of government as the solution and moving toward a role as partner and collaborator. In brief, data analysis is resulting in more effective and efficient government. Ultimately, these new efficiencies will make it possible for local governments to maintain the quality service delivery standards their citizens expect.

## The Technologies

A multitude of emerging technologies will affect how local governments do business in the future. The three technologies discussed below enjoy widespread use throughout North America and are at the forefront of institutional change in how local governments do business. These technologies are geographic information systems (GIS), constituent relationship management (CRM) and 311 systems, and citizen engagement technologies.

### Geographic Information Systems

The very nature of local government is based on location and place. Local governments exist to provide services to citizens who live in geographic proximity to each other, and nearly all data local governments collect have some spatial elements associated with them. The construction of a new housing development creates new street addresses. The police are called to investigate a traffic accident at a particular intersection. Solid waste collection crews take scheduled routes on certain streets to pick up garbage. A streetlight is out in a given block on a specific street.

GIS technology employs a common framework—geographic location—as a means for analyzing a wide variety of data types. Through the creation of datasets that serve as layers (e.g., streets, land parcels, and topology), local governments can integrate diverse datasets via their geographic locations to analyze specific locations and individual properties within the community. The effect of data layering is somewhat akin to flipping through old anatomy textbook films that show how all the systems in the human body work together, with one layer fitting on top of the next layer (Fleming 2005).

When GIS technology was originated, it required immense computing power that limited the use of the technology primarily to federal agencies and major universities. The technology was also very specialized, used largely by technology aficionados who understood its complex world of vectors, rasters, polygons, lines, and points. Over time, however, use of the technology spread. The computing power needed to run GIS technology has become affordable and widely available, and the software has become more intuitive and simpler for the average person to use.

Historically, local governments used GIS technology in their planning and zoning departments and property assessment offices, departments in which decisions on land use dominate the agenda. The spatial nature of the vast majority of local government data allows jurisdictions not only to document and map what is located within the community's geographic boundaries, but also to create alternative scenarios to test assumptions before decisions are made. As more departments saw the analytic power available through GIS, they adapted it for their own purposes. Nearly all local government departments, no matter what their mission and purpose, can benefit from the analytic power of GIS technology (Fleming 2014). Consider just a few of the types of questions that can be answered using GIS:

- If we locate a new community center in this location, how many children under the age of 5 will live within six blocks of it? how many citizens over the age of 65?
- If we allow this level of density in a residential zone, how will that affect the overall streetscape?
- Where are the current fire stations located within the city? What kind of population base do they serve?
- Where are the grocery stores and farmers' markets located within our community? Where are the households living at or below the poverty level located in relation to those venues?

As use of GIS spread to more departments, the way the technology is implemented and managed within local government changed. Independent, stand-alone, single-department programs have given way to an enterprise approach in which core datasets are developed, harmonized, and made available for use by all service departments across the local government organization. Generally these core GIS datasets are maintained and updated by a central GIS office. Specialized datasets required by specific service departments—for example, crime data for police, park amenities and facilities for parks and recreation, or health statistics for public health—

are built using these core datasets, but they are maintained by personnel within the service department. Although the structure of an enterprise GIS looks and operates differently depending on the structure of the local government in which it resides, the general operating principle is the same, namely, that shared data resources enable local government departments to do more together than one service department could do alone.

An enterprise approach to GIS works especially well for local governments wanting to better manage community and neighborhood change, as many service departments have a role to play in improving community quality of life. Pettit and Kingsley (2011) identify key types of data that most local governments have readily available. These include, for example, basic property characteristics, ownership, property tax status, sales prices, foreclosure filings, and building permits.

In an enterprise GIS, such datasets are stored in a manner that allows all service departments to access them if needed. Regardless of where GIS datasets are stored, all such datasets in an enterprise GIS should be developed according to uniform data standards adopted by the local government. Data standards provide a common understanding of what is contained in each dataset. In some sense, they can be thought of as the equivalent of recipe instructions that allow many cooks to create the same dish (Fleming 2014). Maintaining data standards across the local government organization enables GIS analysts to easily combine different datasets to understand possible impacts of decisions in a matter of minutes. Working with a diverse set of stakeholders—neighborhood groups, community nonprofits, state and federal agencies, academia, and private-sector interest groups, among others—local governments can combine their GIS datasets and layers with data from other stakeholders to tackle virtually any community challenge.

The potential for return on investment in an enterprise GIS has also proved substantial. Babinski (2014) reports that the total capital costs to build the GIS program in King County, Washington, came to $10.6 million in 2001, and the annual costs to maintain, operate, and use the system came to $14.6 million. However, a groundbreaking new return on investment study by Richard O. Zerbe of the University of Washington's Evans School of Public Affairs found that the annual benefits received from the program ranged from 6 to 12 times the annual costs, with a projected net benefit estimated at $87 million for 2010. The largest benefits were found in two county departments: natural resources ($54 million) and transportation ($19 million).

GIS technology provides local governments with data access that, in many cases, was not even conceived of three decades ago. For example, many local governments have become involved in brownfields remediation work as a result of the environmental unknowns associated with former commercial and industrial properties. Depending on the type and extent of environmental contamination discovered on a property, the cleanup costs associated with preparing brownfields sites for redevelopment can be extraordinarily expensive. The financial risks associated with such unknowns prevent many developers from moving forward on projects on brownfields sites, leaving the properties vacant, unproductive, and contributing little to the local government tax base.

GIS technology enables local governments to better analyze the potential financial risks associated with brownfields properties, making them more desirable for redevelopment. The creation of historic datasets documents former uses, as well as what materials and chemicals might have been used on the properties. Aerial imagery can reveal former land uses and potential hotspots on a property. Topological and soil data can be used to indicate the possible extent of environmental contaminants and the underground movement of such contaminants over time. The more information available to help determine what may be discovered (i.e., what, if any, contamination exists; what types of contaminants may be found; and where they may be found on the site), the fewer surprises a property holds, and thus the more confidence developers and investors have to move forward with a potentially challenging redevelopment project.

In creating all these data layers, an inventory of available properties emerges that can be used to market brownfields properties. Walsh and colleagues (Fleming 2014) report that in 2008 the Mayor's Office of Environmental Remediation in New York City sought to develop an Internet-based GIS application that could function as a real estate search engine for developers seeking information on brownfields properties and sites in the city. The new application, called the Searchable Property Environmental E-Database, or SPEED, combines brownfields-related data with other local government datasets such as location of schools, hospitals, truck routes, and public transportation hubs. With this arsenal of information, developers have a clear vision of what the local government hopes to have happen with a property and would support should a project be proposed. SPEED has proved its value, having been visited over 3.7 million times by 1.6 million unique visitors since its release.

Intervention and revitalization of residential neighborhoods can also be aided by GIS technology. As Janes and Davis (2011) explain, the City of Baltimore undertook a massive effort to address its substantial stock of vacant and derelict rowhouses. Called Vacants 2 Value, or V2V, the effort involved the creation of an aggressive and targeted code enforcement effort to drive market-based redevelopment in the city. The premise of the program was that rowhouses that were kept up to code would be more marketable.

To undertake this effort in a fiscally responsible fashion, Baltimore worked with contractors to develop two GIS-based management tools: HousingView and CityView. HousingView, an internal application, contains data on tenancy status, assessments, house sales, public or private ownership, inspection districts, and other housing-related attributes that allow the analysis of block and neighborhood development conditions and opportunities. This application helps determine which city blocks and neighborhoods hold the greatest promise for tapping into market-based forces to help to stabilize the area. Inspectors from the code enforcement office use this application to select appropriate neighborhoods for working with property owners on keeping properties compliant with the housing code. CityView, an internal application with a public-facing component, generates maps with housing-related information related to city services and local assets. This application can be used to identify vacant buildings in neighborhoods and provides information on relevant city services and local assets for individuals doing business in the city.

## Constituent Relationship Management and 311 Systems

A constituent relationship management (CRM) system is a technology tool that enables governments to respond to residents' requests for information and service. A 311 system is the broader, centralized customer service system—the people, processes, and CRM technology that enable action in response to the public's inquiries.

CRM and 311 systems give local governments the ability to track and monitor citizen requests for information and services in nearly real time. Using CRM software, customer service agents can collect detailed information from citizens and assign a reference number to a call. Once the information is logged into the CRM system, generally a work order or service ticket is transferred to the corresponding service department. When work is completed, appropriate notes are made, and the work order is closed out by the service department. Should a citizen call back to check on the status of a request, customer service agents can advise on the response thus far.

CRM and 311 systems not only help local governments to provide an improved customer service experience for their citizens, but they also produce a wealth of new data to help leaders understand what services citizens want and need from their local government. Among the types of data and information available through such a system are

- number and types of information requests,
- number and types of service requests,
- time taken to complete service requests,
- percentage of service requests completed within a targeted time-frame,
- geographic location of service requests, and
- trends in citizen requests over time.

These types of data reflect critical business intelligence that can greatly help elected officials and other community leaders determine what services are most critical to their constituents and aid local government service departments in better managing their day-to-day operations.

CRM and 311 systems are transforming how local governments and citizens interact with each other. According to Goldsmith (2012), the idea of government having all the answers when citizens have questions or need help is evolving, and increasingly citizens are being viewed as partners in developing solutions to community problems. Not only are citizens providing data through CRM and 311 systems, but they can also analyze data and help frame solutions based on their review of neighborhood data.

The Mayor's Action Center, the centralized customer service system for Indianapolis and Marion County, Indiana, generates critical data that are used in the combined city–county government's IndyStat Program. The IndyStat Program, a performance measurement and management program, follows a Six Sigma[2] process to pursue continuous improvement within the organization. Measurement is a critical element in the Six Sigma process for determining success, and teams working on a Six Sigma project must analyze data to first understand the nature of the service being reviewed and later to measure progress made on addressing project challenges.

As an example, the mayor's neighborhood liaisons, government employees who work with neighborhood groups on community issues, reported that citizens wanted illegal dumping situations addressed in a more timely manner. IndyStat decided to take the issue on as a Six Sigma pilot project and brought together a diverse group of stakeholders including, among others, the Mayor's Action Center, the mayor's neighborhood liaisons, the Office of the Corporation Council, and the Departments of Code Enforcement, Public Works, and Health.

The Mayor's Action Center provided data on the number of cases being called in by constituents and the nature of the complaints. These data were mapped by the GIS Department to show where the heaviest volume of complaints came from in the city. The Six Sigma team then worked to establish a clear path for issue abatement. In 2008, the resulting pilot program, Clean Sweep, yielded 110 tons of trash from one heavy-hit neighborhood. This neighborhood served as a pilot target area and provided a baseline for measuring statistical change that will be documented as further improvements are made over time.

## Community Satisfaction Surveys and Other Citizen Engagement Technologies

GIS, 311 and CRM systems, and other performance data provide solid quantitative information, but qualitative information that reflects the attitudes of community residents is also important. Community satisfaction surveys measure how residents perceive their community and reflect the public's attitudes toward services provided by the local government. Surveys have long been a favored methodology for engaging citizens in the work of the community. Such data can help local government leaders who want to develop a strategic plan to guide community development, to gauge public support for new initiatives, or to measure citizen satisfaction with services.

Securing citizen input on local government plans and initiatives is critical for the governance of communities. Decisions that affect the whole of a community should reflect the desires of the residents of that community. In an era when demands on personal time prevent many citizens from attending public meetings and workshops, alternate tools for seeking the public's thoughts and reactions about proposed initiatives need to be considered by local governments. Well-designed web surveys that account for sample size and the representation of the whole community offer one solution. More informal methods for exploring public opinion are provided by a host of new applications for web and mobile devices.

In Mesa, Arizona, the mayor and city council were interested in reaching out to the public using a variety of methods, and they wanted to go beyond the feedback received via a traditional community satisfaction survey. In the search for a solution to obtain more dynamic data and information on citizens' ideas, the city chose UserVoice Feedback™, which allows the collection of feedback on proposed ideas that have been prioritized after online voting by citizens. The city launched iMesa, a grassroots citizen investment and improvement effort, to develop community projects to "build a better Mesa." All the ideas posted on iMesa are open for public viewing and debate. Since the introduction of iMesa in 2011, hundreds of ideas have filtered into the city through the website. The city collects all the ideas and schedules a series of volunteer community meetings to discuss and refine the ideas before presenting them to the city council. The mayor and the council work with the city manager to determine how the ideas will be implemented. A significant number of the ideas submitted are expected to transform how the city and the community work together. "If I have 1,000 ideas going into a bucket and one percent of those are great ideas, then I am successful," said Alex Deshuk, Mesa's manager of technology and innovation, who oversees the effort.[3]

## Using Local Government Data

New technologies such as these have helped to generate a wealth of new data and led to what is known as the "big data" movement. In the past two decades, society has moved from discussing data in terms of kilobytes and megabytes to gigabytes and petabytes, and ultimately yottabytes (Manyika et al. 2011). Shark and Cable (2011) note that although the private sector has begun mining business intelligence from big data, using them to identify trends and tackle new business problems, local government leaders have only just begun to explore the possibilities offered.

A central challenge for local governments in using big data stems from the need for personnel who understand how to make the best use of the data. In its predictions for the information technology field for

2012 and beyond, Gartner (2011) suggests that information technology departments will evolve from providing technical support to coordinating information technology activities. This move means that technology users will have a hand in developing solutions to the challenges they face on the job rather than accepting solutions provided by the information technology department. Referred to as *data scientists,* this new breed of local government workers will need to understand both the business questions that need to be addressed using the data and the technology used in analyzing the data. These skills may reside in an individual or, perhaps more importantly, be present in a team brought together to address a particular challenge.

Local governments stand to benefit significantly from more effective use of big data in three key areas: transparency and accountability, performance measurement and management, and innovations in service delivery. These big data issues are discussed below.

## Transparency and Accountability

In recent years, there has been a tremendous call for greater transparency and accountability in government operations as a result of gross mismanagement (Bell, California) and political corruption (Detroit, Michigan). Data have a critical role to play in providing greater oversight of local government operations and management. Citizens, who pay for local government services with their property tax dollars, deserve to have easy access to data to understand how public funds are spent and what results are being received from those expenditures. Web applications that allow an individual to drill down through data to understand what projects are under way and what the associated costs of those projects are help to create much-needed trust between citizens and their local government.

Elected officials and other government executives must regularly make decisions about which public sector programs they should fund, at what levels, and for how long. All too often, policy makers have little or no impartial evidence on which to base their investment decisions. Investments are often made in untested programs that are delivered with little consistency or quality control, and without effective evaluation to determine their effectiveness. Having data offers a solid basis for discussion about program priorities, funding needs, and a host of other daily decisions that make an organization run.

Although all people are equal under the eyes of the law, the services they require from their local government are not. Simply put, equitable service delivery does not necessarily equate to equal service delivery. Local governments that analyze data to better understand the socioeconomic and demographic makeup of their community's population can provide better services to all residents. The amenities offered at a neighborhood park, for example, may vary widely based on the nature of those living near the facility. If a neighborhood's population is composed of empty nesters and few young children, installing playground equipment would make little sense. Likewise, the establishment of public transportation routes and schedules needs to consider where people live and where they work.

## Data-Driven Performance Programs

Performance management programs and process improvement efforts have sprung up in communities across the United States, all designed to provide basic components for data-driven decision making. In 1999, Baltimore instituted CitiStat, the first local government performance management program in which city departments and agencies routinely track metrics to look for substandard performance and propose solutions if service problems are detected.

Behn (2008) employed the global term *PerformanceStat* to describe these types of performance analysis programs, which use data to measure the performance of local government departments in their delivery of programs and services and in operational decisions based on that data. These programs create benchmarks (i.e., the level of service departments can deliver given the necessary staff and resources under normal circumstances), review historical trend data, and ask questions: If the service department isn't delivering its programs and services at the expected level, why not? These programs offer a degree of transparency and accountability that demonstrate public tax dollars are being used appropriately to deliver the programs and services to citizens. The programs are driven in large part by the old adage that what gets measured gets done.

Behn (2008) cautions a PerformanceStat program can easily become the latest government fad if not implemented with integrity and consistency and defines seven common errors governments make in attempting to replicate successes achieved by other governments.

Such programs require strong leadership with a shared vision of the program's purpose, clear responsibilities, and consistent tracking of results. Ideally, PerformanceStat programs should be run in conjunction with some type of continuous improvement program, such as Six Sigma, LEAN, or Total Quality Management, in which the overriding goals are to achieve new levels of excellence, efficiency, and effectiveness in government operations and eliminate errors and waste. Programs designed solely to uncover poor performers tend to stifle creativity.

Another approach to data-driven performance is offered by the International City/County Management Association's Center for Performance Analytics (ICMA Analytics). The mission of ICMA's Center for Performance Analytics is to demonstrate the benefits of professional local government by providing the analytical tools, training, and technical assistance to assist communities in achieving higher levels of performance. In addition, the center disseminates research and best practices.

The center's performance management and analytics software, ICMA Insights™, assists cities and counties in the United States with the reporting, analysis, and application of performance data. The ICMA Insights platform was built in partnership with SAS, a leading company providing business analytics software and services.

The center builds on work begun in 1994 by the Comparative Performance Measurement Consortium, a group of 44 cities and counties whose managers identified a need for accurate, fair, and comparable data about the quality and efficiency of service delivery to their citizens. Consortium members asked the ICMA to coordinate their work and then undertook the challenges of narrowing the choices of services to be measured, identifying desired outcomes of service delivery, defining indicators, and collecting data.

One of the chief values offered by the ICMA Analytics approach is the comparative performance data it provides members. Understanding the nuances and differences in operation is, of course, critical in reviewing such comparative data. To achieve this understanding, participants can choose to compare their data to the full dataset or can use the summary statistics, predefined reports, or interactive reporting tools to customize their comparisons by population, climate, urban density, method of service provision, community demographics, and other characteristics. To facilitate relevant and effective analysis, ICMA streamlined the data collection burden from over 5,000 discrete performance metrics to 900. ICMA Insights is organized in seven service area clusters (Community Attributes, Culture and Leisure, Internal Services, Neighborhood Services, Public Works, Public Safety, Social Services) that look at 18 key sub-areas.

## Innovations in Service Delivery

Historically, local government service departments have tended to operate in a vacuum, focusing exclusively on delivering their own programs and services. Using GIS technology, 311 and CRM data, and other pertinent information, departments can now place relevant data—such as number of graffiti remediation requests, number of abandoned cars, number of code violations, and number of public health complaints—on maps to spot trends in neighborhoods. By examining what is happening in a neighborhood in its entirety, service departments can work together to provide early intervention to troubled neighborhoods. They can also bring in neighborhood groups and community nonprofits to identify other efforts that might be taken to keep neighborhoods healthy and stable. Rather than functioning simply as service providers, local governments are transforming into conveners and collaborators in finding solutions to community problems, using technology to provide a framework for understanding the nature of the problem.

In some cases, the simple act of moving local government data online can improve service to citizens and realize new efficiencies for local government. When the City of Saco, Maine, undertook a major update of its website in 2001, it opted to make as many of the city's business transactions web-enabled as possible. The idea behind the website update was to have constituents "online, not in line," according to Dan Sanborn, city assessor. He estimates that moving records online has reduced the number of citizen inquiries to his office by 2,500 to 3,000 calls annually, thereby enabling him and his staff to redirect their time to other projects (Fleming, Shick, and Stern 2009).

Osborne and Hutchinson (2004) point out that government managers need to actually use data in order to improve their operations and that regular meetings with executive-level leaders provide an important incentive for managers to learn how to analyze and use data. The need for more oversight should not overshadow the drive for continuous improvement or the desire to find creative approaches to achieve new efficiencies and return on investment in local government. Efforts that use local government data should be viewed as an opportunity to create a unique "brain trust" for resolving problems and improving service delivery, not just as a means for ferreting out poor performers. When Minneapolis 311 mapped how service requests for exterior nuisance complaints broke down by districts within the city, they discovered that one of the four supervisor districts generated nearly 33 percent of all exterior nuisance service requests, while another district generated only about 16 percent. Yet both district offices had the same number of support personnel. Given that the demand was so much higher in one district than the other, the city opted to reexamine the allocation of resources within the city's regulatory services department (Fleming 2008).

## Future Developments

The proliferation of smartphones, tablets, and other mobile devices has made it possible for citizens to become partners with their local government in identifying problems and opportunities within the jurisdiction. Mobile applications allow citizens to take a picture, capture GPS coordinates, and submit a service request in a matter of minutes. During the cleanup of the Gulf oil spill in 2010, environmental groups used mobile applications extensively to track where wind and water currents were transporting oil residue along the coastline and identify where cleanup efforts needed to be organized. Known as crowdsourcing, this type of large-group data collection effort can play a vital role in identifying where resources need to be allocated.

However, the effectiveness of crowdsourcing efforts depends largely on the ability to format different datasets so they can be easily integrated and used in conjunction with other datasets. The emergence of big data has led to a related movement toward open data and the push for governments to make their data available in a format anyone can use. The need for easy access to local government data for these applications seems apparent. Manyika and colleagues (2011) note, however, that there is reluctance by many in the public sector to facilitate such access: "The mind-set of employees can be as great a barrier to realizing value from big data as technological inadequacy." Privacy and security concerns have led to the establishment of policy or legal restrictions that prevent data sharing.

Institutional change comes with challenges—implementation costs, reluctant employees, and the general chaos that comes when new business processes are tested and refined—but local governments have a vested interest in taking advantage of the business intelligence provided by new technologies and the data they can analyze. In an age when citizens have a high degree of distrust of government, data can demonstrate how public funds are being used and what the results of those expenditures are. The analysis of data can help elected officials make better decisions and communicate to citizens why those decisions have been made. More importantly, data analysis can lead the way to improvements in business processes that result in local governments being able to maintain high-quality service delivery, as well as cost savings and new efficiencies.

## NOTES

1. "Data and Transparency: Of Governments and Geeks," *The Economist,* February 4, 2010.

2. Six Sigma is a strategy for improving the quality of outcomes by early identification and removal of possible causes of errors or other problems.

3. Alex Deshuk, personal interview, January 26, 2012.

## REFERENCES

Babinski, Greg. 2014. "Show Me the Money: What Was King County's ROI from 20 Years of Using GIS?" In *The GIS Guide for Elected Officials,* edited by Cory Fleming. Redlands, CA: Esri Press.

Behn, Robert D. 2008. "The Seven Big Errors of PerformanceStat." Rappaport Institute/Taubman Center policy brief. Cambridge, MA: John F. Kennedy School of Government, Harvard University.

Fleming, Cory, ed. 2014. *The GIS Guide for Elected Officials.* Redlands, CA: Esri Press.

———. 2005. *The GIS Guide for Local Government Officials.* Redlands, CA: Esri Press.

Fleming, Cory. 2008. "Minneapolis 311 System." In *Call 311: Connecting Citizens to Local Government Case Study Series* (fourth report). Washington, DC: International City/County Management Association.

Fleming, Cory, Robert Shick, and Spencer Stern. 2009. "Saco Information Center and Online Request System." In *Call 311: Linking 311/CRM Data to Performance Measurement and Budgeting Case Study Series* (second report). Washington, DC: International City/County Management Association.

Gartner, Inc. 2011. Gartner Reveals Top Predictions for IT Organizations and Users for 2012 and Beyond.

Goldsmith, Stephen. 2012. "311's Coming Transformations." *Governing,* April 18.

Janes, Ellen, and Sandra Davis. 2011. "Vacants to Value: Baltimore's Market-Based Approach to Vacant Property Redevelopment." In *Putting Data to Work: Data-Driven Approaches to Strengthening Neighborhoods.* Washington, DC: Board of Governors of the Federal Reserve System.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity.* Washington, DC: McKinsey Global Institute, McKinsey & Company Inc.

Osborne, David, and Peter Hutchinson. 2004. *The Price of Government: Getting the Results We Need in an Age of Permanent Fiscal Crisis.* Washington, DC: Basic Books.

Pettit, Kathryn L. S., and G. Thomas Kingsley. 2011. "Framework: The New Potential for Data in Managing Neighborhood Change." In *Putting Data to Work: Data-Driven Approaches to Strengthening Neighborhoods.* Washington, DC: Board of Governors of the Federal Reserve System.

Shark, Alan R., and Susan Cable. 2011. *Civic Media in Action: Emerging Trends and Practices.* Alexandria, VA: Public Technology Institute.

ESSAY

# From Tornadoes to Transit

## *How Data Sharing Has Strengthened Community Development in the Twin Cities*

Jacob Wascalus and Jeff Matson

Currently, the Twin Cities metropolitan region in Minnesota enjoys an active data-sharing environment that plays an increasingly prominent role in neighborhood and community development. Many people and institutions populate this milieu, including the county and municipal departments that collect, generate, and disseminate community data and the nonprofit intermediaries that study and apply these data. Situated in the middle, of course, is the public—the community activists and neighborhood groups that, to improve their surroundings, partner with and rely on the active involvement of government departments and intermediaries. But although community residents have always sought the betterment of their environments, the ingredients necessary for successful community change have taken time to develop. Similarly, the ability of the local government and nonprofit agencies to respond to pressing immediate needs has evolved.

In May 2011, a tornado with winds topping 110 miles per hour tore through North Minneapolis, one of the poorest neighborhoods in the Twin Cities area, and left a four-mile-long path of destruction in its wake. The twister damaged nearly 1,900 properties—274 of them substantially—and left hundreds of people homeless.[1]

In the aftermath of the storm, officials from the City of Minneapolis, several neighborhood associations, and various nonprofits turned to the University of Minnesota's Center for Urban and Regional Affairs (CURA) to act as the de facto data cruncher and information distributor.

CURA, the lead local partner in the National Neighborhood Indicators Partnership (NNIP) in the Twin Cities area, quickly got to work providing assistance in the tornado response effort. It created and distributed maps that helped responders identify the needs of the community (such as where food, clean water, and clothing were needed), the severity of each house's damage, and the locations of suspected vacant properties and rental housing.

CURA's assistance was vital to the response effort because few, if any, organizations or government departments were adequately positioned to pull together the information and data needed for the post-tornado response. The urgent circumstances of this natural disaster required immediate action, but CURA's help would not have been possible without the presence of an active data-sharing environment that had developed in the Twin Cities over the preceding decades. In fact, if the tornado had struck 15 years earlier, CURA's help would have largely been limited to on-the-ground support.

This essay traces the history and current status of the data-sharing environment in the Twin Cities,[2] with a particular emphasis on the current landscape of nonprofit intermediaries, like CURA, that apply community data to improve neighborhoods throughout the Twin Cities metropolitan region. What emerges is a picture of a data-sharing environment that developed organically, starting with key decisions made by state-level government entities and continuing to the data-oriented projects that nonprofit intermediaries undertake to address community needs.

## Taking Root in State Initiatives

Community data users in the Twin Cities metropolitan region may take for granted the abundance of data at their disposal, but access to community information has not always been as easy as downloading a ready-to-use geographic information system (GIS) file or opening up a regularly distributed dataset. In fact, the current data-sharing environment in the Twin Cities owes much to actions undertaken decades ago by a handful of state and regional government agencies and professional associations. Collectively, their actions achieved three important and necessary steps in shaping the current data-sharing atmosphere: establishing important standards for recording and annotating data, forging the interagency and professional connections needed both to exchange ideas and to advance the overall data-sharing environment in the state, and developing innovative platforms for distributing data.

Efforts to standardize data got under way with the Land Management Information Center, a state agency created in 1977 that was devoted to coordinating and providing GIS services within Minnesota's state government. The goal was to establish a central GIS "shop" for the entire state. The Land Management Information Center worked with multiple government agencies that generated a variety of data, which were often recorded and described without much thought to inter-department uniformity. Recognizing the benefits of adapting a consistent annotation system, the center established state metadata guidelines for describing key aspects of these state-generated datasets, such as data quality information, distribution information, and so on.[3]

Further steps for standardization were undertaken by the Governor's Council on Geographic Information, an advisory group created by executive order in 1991 to guide and work on statewide GIS policy (with administrative and technical GIS services still being provided by the Land Management Information Center). A committee from the council adopted a process for establishing additional state standards, and even though the standards in question were not applicable to local governments, representatives from local municipalities served on this committee. The idea was that city and county departments, having helped form state-level standardization guidelines, might adopt these standards for their own departments (see Craig, Baker, and Yaeger 1996).

The Council on Geographic Information also proved to be a useful forum to network and exchange ideas. Composed of 18 representatives from a cross section of sectors—including federal, state, and local governments, as well as higher education and the private sector—the council convened the state's top GIS experts to shape GIS policy and direction (Craig 2005). One of the council's guiding principles was to "promote geographic information as a public resource that should be widely shared with and available to interested parties."[4] It should come as little surprise, then, that a product of the council's work was the creation of the Minnesota Geographic Data Clearinghouse, an online, searchable repository of geographic data that provided (and continues to provide) access to hundreds of datasets developed and maintained by state and local governments.[5] Launched in 1997, the Geographic Data Clearinghouse was the first Minnesota portal that the public could use to link to these datasets (Minnesota Governor's Council on Geographic Information 1997).

## Data Sharing Emerges Locally

In the Twin Cities, the Metropolitan Council, the regional planning agency that provides a broad range of services to the 182 cities and municipalities across the seven-county metropolitan region, launched its own regional GIS network (MetroGIS) and data-sharing portal (DataFinder). Since its inception in 1995, MetroGIS has sought to "institutionalize the sharing of accurate and reliable geospatial data and information" to help organizations and local government departments carry out their work more effectively. In pursuit of this goal, in 1998 MetroGIS launched DataFinder, an online GIS data clearinghouse that is similar to the Minnesota Geographic Data Clearinghouse but focuses on regional and local geographies. The site relies on the input and direction of hundreds of representatives from agencies and organizations across the metro area—people from all levels of government, including city and county departments, watershed districts, school districts, and utility companies; private and nonprofit organizations; research organizations; and private citizens. In fact, many of these people contribute the data that constitute the portal's catalog of datasets.

For people interested in community development, DataFinder provides the mother lode of community information: more than 275 datasets on everything from administrative and political boundaries to outdoor recreation infrastructure. Moreover, the vast majority of the data is available to the public and instantly downloadable.

But the key dataset—the foundation on which most community-based geospatial analysis is conducted—is the MetroGIS Regional Parcel Dataset. Since 2002, each of the seven counties in the Metropolitan Council's geographic area has created and updated, quarterly, an ArcGIS shapefile containing the complete record for every parcel in its boundaries. Each of the approximately 900,000 records contains a unique property identification number and comprehensive information on 55 property attributes, including lot size, building size and age, owner name, acquisition price and date, and tax assessment amount.[6] Other geography shapefiles are available through MetroGIS's DataFinder, including different local taxing districts and even census blocks already populated with demographic, housing, and economic data. However, none of these shapefiles provides the granular detail available through the individual parcel files contained in the MetroGIS Regional Parcel Dataset.

In addition to making their data available through MetroGIS's DataFinder, many city and county government offices are beginning to record their data in automated record-keeping systems and distribute these datasets, on request, to government, academic, and qualified nonprofit recipients. For example, each month the Hennepin County Surveyor's office captures sheriff sale[7] data that are published in the legal newspaper of record. The office places the details of each foreclosure in an ArcGIS point file, which is a useful format for conducting spatial analyses in conjunction with the parcel data available through DataFinder.

Another government office, the Minneapolis Assessor's office, evaluates the physical condition of each structure in the city and records this rating in an internal database from which it can create and distribute, on request, an Excel file that links each building condition to a property's identification number. These ratings can easily be joined with the data from the Hennepin County parcels. Users then have the ability, for instance, to determine if owner-occupied properties have better condition ratings than rentals or to examine the relationship between sales price and condition.

Other examples of information that is being digitized and shared by city and county departments include a list of building permits in Saint Paul and Minneapolis, a registry of vacant and boarded buildings, and a database of housing code violations.

## The Role of Intermediaries

Over the past decade, a few Twin Cities–based nonprofit intermediaries have taken advantage of this emerging movement of shared county and municipal information to launch a range of data-oriented projects that help facilitate informed community change. These organizations include CURA, HousingLink, and Wilder Research.

### Center for Urban and Regional Affairs

CURA was founded in 1968 following the civil unrest and rioting that had roiled urban areas across the United States. Its purpose was to make the university relevant and responsive to the concerns of local residents and to act as the school's community-engagement arm. Today, CURA pursues this same mission by connecting university resources—student researchers, academics, course assignments—with people and organizations in urban and rural communities to collaboratively study and address issues on a broad range of topics, from urban housing and economic development to rural food systems and the experience of immigrant populations in small towns.

Over its 45-year history, CURA has established a reputation for being an integral centerpiece and facilitator of collaborative community change. Moreover, because of its academic setting, CURA is able to apply the latest research techniques and technologies to its projects. A by-product of this approach has been the steady advancement of innovative community development tools—including data-oriented tools—available to the people and groups who are working to improve their communities. Two projects, the Minneapolis Neighborhood Information System (MNIS) and Minnesota-3D (M3D), illustrate CURA's important role in the data-oriented community development environment in the Twin Cities and demonstrate the organic process by which this environment has grown.

## Minneapolis Neighborhood Information System

Although innovation in research methods and technology applications has been a regular feature of CURA's work, one project in particular, MNIS, solidified its data-oriented program offerings while significantly advancing the broader data-sharing and application environment in the Twin Cities. Initiated in 2001, the MNIS project sought to identify specific properties in Minneapolis that were at risk of foreclosure or abandonment.

At the time, Minneapolis had been experiencing a net loss of housing stock, with more demolitions occurring than new construction. One neighborhood in particular—the Central neighborhood, in South Minneapolis—experienced this phenomenon to such a degree that one of every six houses was vacant or had been demolished. Recognizing housing contraction as a sign of instability and disinvestment, the Central Neighborhood Association worked with CURA to create an early warning database by combining a variety of municipal datasets that each represented different indicators of abandonment. The datasets, which CURA obtained after working with various City of Minneapolis departments, included property tax delinquency, water billing arrears, poor building condition, nonowner occupancy, proximity to abandoned properties, and proximity to high-crime areas.[8]

Although MNIS was initially envisioned as a neighborhood-specific database, it grew to include property information for the entire city and was ultimately housed and maintained by the city government. CURA, which was at the center of this expansion, eventually served as the coordinating force for 65 neighborhood groups and various departments at the City of Minneapolis. The city maintained the MNIS site until 2007, when the city upgraded its GIS platform. The updated parcel data files are still available to CURA and are used in community planning and policy analysis.

In addition to the successful application of municipal data to identify properties that were at risk of foreclosure and abandonment, several key developments occurred because of this project. First, the foundation of data on which MNIS was based was made readily available to the public for the first time. These data included information from the Minneapolis Assessor's office (e.g., parcel data such as estimated market value, property condition, owner occupancy, absentee landlordism), the Minneapolis Inspections office (e.g., code violations), the Minneapolis Regulatory Services office (e.g., problem property inventory, vacant and boarded buildings), the Minneapolis Utilities Billing Department (e.g., delinquent water bills), and Hennepin County Taxpayer Services (e.g., property tax delinquency).

Second, the MNIS project helped community organizations recognize the power of GIS and the multiple uses of administrative data beyond simply identifying problem properties. For example, a key by-product of the project was community capacity building. CURA organized regular GIS trainings for neighborhood organization staff, and the GIS users from these organizations participated in regular meetings to exchange ideas and share examples of how they were applying the administrative data to answer questions about their respective communities.

Third, CURA formally established Community Geographic Information Systems, a program dedicated to providing GIS services for neighborhood groups. This program has since operated as a community-oriented walk-in technical assistance center with the overall goal of improving access to and usability of community-level geographic data through mapping, data analysis, and web-based applications. Community Geographic Information Systems has offered data and mapping assistance to scores of community organizations on a wide range of community-related data projects. After the founding of this program, CURA was invited to join NNIP.[9]

Having firmly established its neighborhood-level GIS bona fides with the MNIS project, CURA has since collaborated on hundreds of community data projects and continues to broaden the data-sharing environment in the Twin Cities region. An additional effort—the M3D project—underscores CURA's role as a proponent of expanding the use of GIS tools and incorporating previously unreleased or inaccessible government data holdings into its work.

## Minnesota-3D

When M3D was initiated in 2004, the Twin Cities region had been experiencing a widening spatial mismatch between where people lived and where people worked, a problem particularly pronounced in areas where populations of color resided. CURA recognized this problem and partnered with several government and nonprofit agencies to create M3D, a web-based GIS assessment tool that provides a comprehensive snapshot of the region's housing and labor markets, commuting patterns, transportation networks, affordable-housing locations, and development opportunities.[10] CURA and its partners created this application so that planners, developers, businesses, and the public could have easy access to the information they needed to make informed policy, investment, and infrastructure decisions.

The M3D project stands out for two reasons. First, the database synthesized a huge catalog of datasets—more than 90, in fact, from sources ranging from regional and municipal agencies to nonprofits. Second, it empowered a wide spectrum of users to access this trove of information with nothing more than a personal computer and an Internet connection.

M3D doesn't simply incorporate the foundational demographic information available through the US Census; it ties that information to literally scores of other datasets—data that CURA, from its centralized role in the project, helped acquire. To illustrate the benefits of using this tool, take the relatively simple example of a neighborhood organization in Saint Paul seeking to enhance transportation options for its residents. The neighborhood could quickly map the work locations of its residents, overlay bus routes, and see at a glance where transit service was lacking.

This information could be used to advocate for new or increased service through that neighborhood. In the suburbs, communities could also use M3D to foster collaboration with other cities in transportation planning. For instance, the City of Chaska, which is approximately 30 miles from downtown Minneapolis, worked with neighboring communities Eden Prairie, Chanhassen, and Bloomington on roadway improvements and succeeded in increasing transit service after using M3D to demonstrate that residents from each city traveled through neighboring communities on their daily commutes. The M3D site also features employer and employee dynamics, including home–work locational information and the demographic characteristics of an area's workers and residents. This information could be used to examine the industries located in a neighborhood or in which an area's residents were employed. It could also be used to attract businesses that could employ residents of a specific area or establish training programs that matched skills demanded by local jobs.

Several of the more than 90 datasets incorporated into M3D had not previously been available to the public: Local Employment Dynamics data from the Census Bureau; Minnesota Department of Revenue sales tax and business data; and the affordable-housing inventory maintained by HousingLink, a government-created nonprofit. (For more on HousingLink, see below.) CURA continued to update data on the M3D site until early 2010.

To access M3D, users need only a personal computer and an Internet connection—no specialized desktop GIS software is necessary, nor is there a need to spend a significant amount of time learning how to use the interface. CURA and the other project contributors chose to create the application in this way to ensure that anyone with the inclination to learn about his or her community could do so easily, without restriction. For example, by using M3D, city planners could evaluate future demand for transit service, community developers could identify markets for infill development, and community groups could gather information on subjects important to them in order to lobby decisionmakers.

## HousingLink

CURA's role in applying community data through its Community Geographic Information Systems program has enhanced its reputation as an effective facilitator of community change. But another nonprofit intermediary pivotal to the Twin Cities' community development world—HousingLink—is valuable in its role as a data collector and distributor.

Since its inception in 1997, HousingLink has served as Minnesota's main hub for information and data related to affordable housing. It was formed as a result of a 1995 lawsuit and tasked with becoming the centralized resource for Section 8 voucher holders and people living in public housing to better understand their housing options.[12] To meet this need, HousingLink, over the past 16 years, has tracked affordable rental housing across the state (the majority of which is in the Twin Cities region) and makes the information available through two interactive, web-based mapping applications called hList and Streams.

Through hList, HousingLink tracks the availability of public and private affordable-housing units across the state and publishes actual rental listings where there is a vacancy or an open waiting list. To do this, HousingLink offers private-market landlords and other housing providers the opportunity to list, for free, their properties directly on the hList web application. It also regularly draws additional listings from a variety of public sources, both print and online. This comprehensive approach provides visibility to more than 15,000 listings every quarter.

Streams, HousingLink's online database of rent-subsidized properties, provides detailed information about each dwelling's location, government funding source (or "stream"), affordability commitments, and renter income qualifications. Although this site is available to the public, it is intended primarily for researchers, advocates, and policymakers—anyone interested in the supply of subsidized housing. HousingLink gathers this information by working with a range of entities that provide housing subsidies, including federal agencies like the US Department of Housing and Urban Development and the US Department of Agriculture, state and local entities like Minnesota Housing (the state's affordable-housing agency), Minneapolis's Community Planning and Economic Development Department, and a variety of other local government and nonprofit funders.

As a testament to their value, each month more than 30,000 people visit the hList and Streams web applications. Although HousingLink itself doesn't participate in community development projects (it serves only as a data clearinghouse and resource center), many people and institutions use the information available on its website to inform their projects. For instance, the Beacon Interfaith Housing Collaborative, a supportive-housing organization with multiple shelters and housing developments throughout the Twin Cities, uses HousingLink's Streams application to gather intelligence on potential future housing projects.

By identifying areas with a dearth of subsidized housing options, the organization can then reach out to faith-based institutions in each area to explore potential partnership possibilities. In addition, because of the breadth of its hList database, HousingLink publishes *Rental Revue,* a quarterly subscription-based rental housing report that reflects the local pricing levels of the Twin Cities' private rental market for 65 cities within the seven-county region. CURA has used this information to create detailed reports that highlight affordability and track gentrification at the neighborhood- and census tract–level of geography.

## Wilder Research

Another Twin Cities–based intermediary is Wilder Research, the nonprofit research arm of the Saint Paul–based Amherst H. Wilder Foundation. Since issuing its first study in 1917, Wilder Research has supported community-building efforts by conducting practical and useful research for its parent foundation, as well as an array of nonprofit and government agencies, service providers, and policymakers, all of which use the organization's research to better understand the issues that affect the respective communities they work in.

Wilder Research is innovative in its approach to studying community issues and regularly applies a range of data to its research projects. But what sets this research organization apart from other nonprofit intermediaries is its work in applying data the organization itself collects into broader syntheses of community data.

One project in particular is a good illustration of Wilder Research's use of primary and secondary data: its baseline indicators report for the Minneapolis–Saint Paul Central Corridor light-rail transit line.

In 2010, the Metropolitan Council, the Twin Cities' regional planning authority, broke ground on an 11-mile light-rail transit line that will link downtown Minneapolis to downtown Saint Paul. The line will run through the University of Minnesota campus and along a major commercial corridor that abuts more than a dozen neighborhoods. Although the light-rail transit line will certainly enhance mobility for many along its corridor, its construction has sparked controversy over the effects on the low- to moderate-income neighborhoods that are situated along the route. To gain a better understanding of how the construction of the light-rail transit line and its eventual operation will change the lives of the people and businesses that reside or operate near it, Wilder Research, on behalf of 13 foundations that fund projects in the area, developed a baseline indicators report on the conditions along the corridor. To develop these indicators, Wilder Research applied a range of datasets from federal, state, and local sources while incorporating information it generated itself through household surveys. The resulting Central Corridor Tracker, which will be updated annually, is an innovative synthesis of data types that is a hallmark of other Wilder Research work.

## Trends in Applications

Over the years, CURA, HousingLink, and Wilder Research have each worked to improve communities across the Twin Cities through their use and dissemination of community data. A hallmark of their work has been continual efforts to innovate in their application of these data: CURA through its Community Geographic Information Systems program, HousingLink through its hList and Streams web applications, and Wilder Research through its community research projects. The emphasis on innovation will continue. Several recent trends in the dissemination, presentation, and application of community information provide a glimpse of the future direction of data-oriented community development.

People who wish to connect directly to a data provider's servers to import and view spatial data can now do so through new data-access options called map services, which obviate the need to download, store, and update datasets. With map services, the entities providing the data store the data files themselves and are responsible for updating them. Users of these services can access the data through ArcMap (both the desktop version and the online version) and open source GIS software such as Mapserver. Two local entities in the Twin Cities area are currently disseminating spatial data through this option: MetroGIS's DataFinder (discussed above) and Ramsey County, the home of Saint Paul. Some current examples of map services include aerial photographs (with files that are typically massive) and simple shapefiles for roads, land use designations, and political boundaries.

In addition to providing direct links to spatial data files, some entities have begun disseminating and presenting public information through online web-mapping applications, such as ArcGIS Online. Although users cannot access the underlying data, many of the mapping services being provided by these online tools are highly customizable and can be embedded into other websites, blogs, and social media, making them not just easily sharable but accessible to everyone. CURA maintains one such application. Through it, users can view characteristics of housing and other socioeconomic variables by neighborhood in both Minneapolis and Saint Paul. Most of the metropolitan counties maintain web-mapping applications to disseminate public information, such as the legal characteristics of properties within their boundaries and the locations of recent crimes. Some of these counties are even creating applications designed for mobile phones. Carver County's phone app, for example, provides on-the-go users with information related to trail locations and other active-living events in that area.[13]

Finally, to draw on the collective intelligence of a cross section of community-oriented problem solvers, local governments in the Twin Cities region have begun to host collaborative problem-solving events called hack-a-thons and code challenges. These events bring together a diverse group of community members (e.g., programmers, data providers, and community activists) and charge them with working toward the common goal of creating useful applications, visualizations, and tools for public consumption. To heighten the likelihood of producing something worthwhile, these activities are compressed events that transpire over a short period of time, such as a weekend or a long day (or even just eight hours). These events underscore the value of making use of public data. As a case in point, various sponsors, including the Minneapolis-based McKnight Foundation, helped CURA host Visualizing Neighborhoods, an eight-hour hackathon that convened data visualizers, designers, artists, scientists, civil servants, and neighborhood leaders to "explore how data can be used for research, analysis, mapping, outreach, engagement, and communication" in neighborhoods throughout the Twin Cities. The goal of the event was to initiate dialogue, strengthen communities, experiment with data, and design prototype projects for neighborhoods. Community data were central to this entire process.[14]

## Remaining Barriers

From the formation of the Land Management Information Center more than 35 years ago to the web applications and hack-a-thons debuting today, the data-sharing environment in the Twin Cities continues to advance rapidly. Still, despite the seemingly open data policies of some municipal departments, the public continues to face often byzantine processes to obtain community data, and much of the time it cannot access this information at all. In fact, most of the important datasets for community work, such as the GIS parcel shapefiles for the seven counties in the metropolitan region, are largely restricted to government and university workers. Other data may technically be available to the public but are not accessible (i.e., downloadable in raw form) for manipulation, such as Minneapolis crime data, which are presented only on maps in PDF form.

As consumers of data and maps have grown more sophisticated, they have demanded more of the information, applications, and tools fundamental to effective community building. The challenge over the next decade will be to continue to foster the creative collaborations that bring data producers and intermediaries together with the communities that have questions to answer or problems to solve. Unfortunately, the need to address community questions and problems will be a reality for a long time. But if the history in the Twin Cities is any indication, the data and the people to apply it will be there, too.

## NOTES

1. The postdamage assessments were conducted by the Federal Emergency Management Agency.

2. For the purposes of this essay, the Twin Cities metropolitan region includes all the communities in the seven counties (Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington) encompassing and surrounding Minneapolis and Saint Paul.

3. The state metadata guidelines were based on the federal geographic metadata guidelines. See "Minnesota Geographic Metadata Guidelines, version 1.2," Minnesota Governor's Council on Geographic Information, http://www.gis.state.mn.us/pdf/Mgmg1_2.pdf for more information. See also Craig (2005).

4. To see other guiding principles of the council, visit "About the Council," Minnesota Governor's Council on Geographic Information, http://www.gis.state.mn.us/about.htm.

5. In 2009, the Land Management Information Center and the Governor's Council were replaced by MnGeo, a state agency legislatively chartered to coordinate GIS within Minnesota. MnGeo houses the Minnesota Geographic Data Clearinghouse, for which GeoGateway is the search engine.

6. For a complete list of the 55 attributes, see www.datafinder.org/metadata/MetroGIS_Regional_Parcels_Attributes.pdf.

7. A sheriff sale is the public auction of a mortgaged property that is undergoing foreclosure.

8. For more information on MNIS, see Mardock (1998), Goetz and Schaffer (2004), Matson (2004) and Matson and Nelson (2004).

9. CURA joined NNIP in 2007.

10. M3D is discussed in Matson, Nelson, and Mahlik (2007).

11. As of September 2014, the M3D website is still online with the historic data at http://map.deed.state.mn.us/m3d/index.html. In addition, some of the site's features, such as the ability to create neighborhood-level reports and commutesheds, are now available in the Census Bureau's On The Map website at http://onthemap.ces.census.gov/

12. Hollman v. Cisneros Consent Decree. For more information, visit http://www.housinglink.org/.

13. See Carver County's online mapping services at http://www.co.carver.mn.us/departments/admin/IS/gis_mapping_applications.asp.

14. For more information about this event, visit "Visualizing Neighborhoods: A Hackathon for Good," CURA, http://www.cura.umn.edu/visualizingneighborhoods.

## REFERENCES

Craig, William J. 2005. "White Knights of Spatial Data Infrastructure: The Role and Motivation of Key Individuals" *URISA Journal* 16 (2): 5–13.

Craig, William J., Michael N. Baker, and Donald P. Yaeger. 1996. "Coordinating Data in Minnesota." *GeoInfo Systems:* 30–32, 34.

Goetz, Edward G., and Brian Schaffer. 2004. An Evaluation of the Minneapolis Neighborhood Information System. *CURA Reporter. 34* (4): 18–23.

Mardock, Lori. 1998. "Predicting Housing Abandonment in Central: Creating an Early Warning System." NPCR 1089. Minneapolis: Neighborhood Planning for Community Revitalization (NPCR), Center for Urban and Regional Affairs, University of Minnesota. http://www.cura.umn.edu/sites/cura.advantagelabs.com/files/publications/NPCR-1089.pdf.

Matson, Jeffrey. 2004. "Neighborhood Planning and GIS: Case Studies from the Minneapolis Neighborhood Information Systems." *CURA Reporter 34* (3): 9–16.

Matson, Jeffrey, and Kris Nelson. 2004. "Minneapolis Neighborhood Information System: A Community, University, and City Collaboration." *CURA Reporter 34* (2): 9–16.

Matson, Jeff, Kris Nelson, and Elissa Mahlik. 2007. "Minnesota 3-D: Bringing Planning and Community Data Together Online." *CURA Reporter* 37 (4): 28–39.

Minnesota Governor's Council on Geographic Information. 1997. "Laying the Foundation for a Geographic Data Clearinghouse." St. Paul: Minnesota Governor's Council on Geographic Information.

# 3
# Progress in Data and Technology

Changes in the quality and accessibility of data and technology over the past two decades have fundamentally shifted the possibilities for the development and use of neighborhood indicators. This chapter gives an overview of the types of data available to create subcity indicators, illustrating important milestones and shifts over the past two decades. It does not aim to be a detailed primer or comprehensively cover all sources of data, but references throughout the chapter will provide additional resources for the reader. The chapter is divided into two parts, data and technology. The data section describes common secondary data sources and methods for collecting primary data for small areas. It then reviews several notable trends that are expanding the applications, accessibility, and sources of neighborhood indicators. The technology section introduces how the advances in hardware, software, and web services have altered how data are produced, processed, shared, and visualized. The chapter concludes with the implications of the progress for nonprofits and governments who are working to improve their communities. Greg Sanders's essay, which accompanies this chapter, touches on many of the topics from a technologist's perspective and suggests several areas of opportunity for incremental and fundamental change.

## Data for Neighborhood Indicators

### Secondary Data Sources

Neighborhood indicators are most commonly derived from secondary data that have been collected for other purposes by government agencies, nonprofit organizations, or private firms. The original uses of the data could be to fulfill legislative or oversight requirements, to administer government programs, or to monitor a specific industry. Using secondary data sharply reduces the time required to create neighborhood indicators compared with primary data collection (see below). Secondary data sources with national coverage also allow for easier comparisons within and across cities and regions. The discussion in this section is organized by the sector (federal government, state and local governments, and commercial firms) that distributes the various data sources. In reality, the data sources overlap: federal government aggregates private-sector data to share; state and local governments submit data to the federal government, which redistributes them as a national series; and private firms repackage federal and local data into new products.

#### *Federal Government Data*

Before the late 1980s, the Decennial Census served as the primary source for neighborhood indicators for all areas in the United States. Although data storage and processing were daunting given the technology of the time, this rich source of data offered indicators on a wide variety of individual, household, and housing characteristics that could be compared across geographic areas. Over the next decade, federal agencies ramped up the production and release of data files at the address or small-geographic level. These data were as diverse as Common Core of Data on public school and student characteristics, Zip Business Patterns on employment and establishments, or A Picture of Subsidized Housing (see table 3.1). When first issued, the data could be ordered as data tapes or CDs. In the 2000s, as use of the Internet became commonplace, agencies transitioned most existing data series to online distribution. The increased power and prevalence of desktop computers expanded the audience for large (by that day's standards) administrative datasets. Even so, use of the online data was gradual. The 1999 review of the data systems of the founding National Neighborhood Indicators Partners

**Table 3.1.** National Datasets with Small-Area Data, September 2014

| Dataset | Smallest geography | Frequency of source data update | Years available electronic format | Description |
|---|---|---|---|---|
| | | | **Multitopic** | |
| Decennial Census | Block | Every 10 years | 1990, 2000, 2010 | Demographic characteristics of people and households, occupancy and tenure of housing units |
| American Community Survey | Block group | Annual | 2005/09–2008/12 | Sample survey with demographic, economic, and housing characteristics of people, households, and housing units |
| | | | **Housing** | |
| Home Mortgage Disclosure Act | Census tract | Annual | 1990–2012 | Home mortgage applications, loans, and loan denial rates by applicants' income, and race/ethnicity. Includes information on high interest rate lending |
| Consolidated Plan Data (HUD Special Tabulations) | Census tract | Annual | 2000, 2005/09, 2007/11 | Numbers of households and housing needs indicators (e.g., housing cost burdens) by HUD income category (e.g., less than 30% of area median income, 30–50% of area median, etc.) |
| Low Income Housing Tax Credit | Point | Annual | 1995–2012 | A variety of project-level data on Low-Income Housing Tax Credit projects including low-income units, project type (new construction/rehabilitation), and number of bedrooms |

*(continued)*

**Table 3.1.** *(Continued)*

| Dataset | Smallest geography | Frequency of source data update | Years available electronic format | Description |
|---|---|---|---|---|
| HUD Neighborhood Watch | Zip Code | Monthly | 2002–2014 | Information on FHA loans that became 90 days delinquent during their first 2 years, by geographic area or originating lender |
| Multifamily Assistance and Section 8 Contracts | Point | Every 2 to 3 months | Since 2001 (point in time) | Project and tract-level data on housing units (total and by bedroom size) covered by Section 8 contracts and expiration year for those contracts |
| A Picture of Subsidized Housing | Point/ census tract | Annual | 1997, 1998, 2000, 2004–2013 | Characteristics and counts of the households in public and subsidized housing units |
| Government-Sponsored Enterprise Mortgage Data | Census tract | Annual | 2000–2007 | Datasets on single-family and multifamily mortgage purchases by Fannie Mae and Freddie Mac. Includes information on the income, race, and gender of borrower as well as the loan-to-value ratios, purchase/refinancing, and affordability |
| HUD U.S.P.S. Address Vacancies data | Census Tract | Quarterly | 2005–2014 | Number of addresses receiving mail and the number not receiving mail by length of vacancy |

| | | | | |
|---|---|---|---|---|
| Location Affordability Index | Block group | Unknown | 2006/2010 | Estimates of the percentage of a family's income dedicated to the combined cost of housing and transportation in a given location for different types of households. Includes the independent variables that were used to construct the index |
| National Preservation Database | Point | Varies by data source | Point-in-time | Integrates information on all housing subsidies for each federally subsidized project |
| | | **Income and employment** | | |
| Federal Individual Income Tax Files (SOI) | Zip Code | Annual | 1998–2008, 2011–2012 | Data on filers' income and age distribution, savings information, EITC and RAL, tax providers |
| Zip Business Patterns | Zip Code | Annual | 1994–2011 | Summary data on numbers of private business establishments and employment by number of employees and type of industry (e.g., manufacturing, trade). Includes establishments in retail service categories such as grocery stores, banks, and dry cleaners |
| Old-Age, Survivors, and Disability Insurance (OASDI) beneficiaries | Zip Code | Annual | 2003–2012 | Includes the number and type of beneficiaries (retired, disabled, surviving family members); the amount of benefits paid; and the number of beneficiaries over age 65 |
| Small Area Income and Poverty Estimates | School District | Annual | 1989–2012 | Number of people in poverty by age group and the median household income |

*(continued)*

**Table 3.1.** *(Continued)*

| Dataset | Smallest geography | Frequency of source data update | Years available electronic format | Description |
|---|---|---|---|---|
| LEHD Origin-Destination Employment Statistics (LODES) | Block | Annual | 2002–2011 | Data on the number and characteristics of workers by location of residence and location of workplace, and commuter flows from block to block |
| | | **Neighborhood assets** | | |
| NCES Common Core of Data | Point | Annual | 1986/87–2012/13 | Includes description of schools and school districts: name, address; data on students and staff, including demographics |
| FDIC-Insured Summary of Deposits | Point | Weekly | 1994–2013 | Includes locations of full-service and limited service bank branches |
| NCCS Core Public Charities Database | Point | Annual | 2000–2012 | Descriptive information from charities' initial registration with financial variables from the Tax Form 990. Only for nonprofits required to file taxes ($25,000 revenue and above) |
| Public Libraries in the US Survey | Point | Annual | 1992–2012 | Provides national descriptive data on the status of public libraries in the United States collected from over 9,000 public library systems with over 17,000 public library outlets |

| | | | | |
|---|---|---|---|---|
| | | **Environment** | | |
| EPA Smart Location Database | Block group | Unknown | 2006–2010 ACS, 2010 LED, various others | A collection of indicators from various sources that help to measure location efficiency. Includes indicators related to housing density, diversity of land use, neighborhood design, destination accessibility, transit service, employment, and demographics |
| EPA Enforcement and Compliance History Online (ECHO) | Point | Varies by data source | Varies by data source | A collection of many individual EPA data sources information on regulated facilities related to air emissions, surface water discharges, hazardous waste, and drinking water systems. Contains permit data, inspection dates and findings, violations, enforcement actions, and penalties assessed |
| Toxics Release Inventory | Point | Varies based on data source | 1997–2014 | Information about hazardous waste sites, facilities that produce and release air pollutants or toxics, and permits to discharge to water |

*Note:* For updated information and web links, see http://www.neighborhoodindicators.org//library/catalog/list-national-data-sets-small-area-data mentioned only the Decennial Census, even though several of the new data series had started by that time (Kingsley 1999). Table 3.1 lists many of the federal small-area datasets. The remainder of this section highlights a few of these data sources that represent major trends or important innovations in data production or dissemination.

The Home Mortgage Disclosure Act (HMDA) is one example of a major advancement in data collection and provision. HMDA was enacted by Congress in 1975 and implemented by Federal Reserve Board regulations. HMDA legislation requires most lenders to report home mortgage applications with loan attributes, applicant characteristics, and the census tract of the property. Beginning in 1992, lenders were required to make the loan-level data public, and the Federal Financial Institutions Examination Council began to produce CDs with an easy-to-use system to query the data (McCoy 2007). Although the HMDA was motivated by the need to monitor financial institutions' investments in communities and identify potentially discriminatory lending patterns, it also enabled the creation of indicators on the race and gender of new borrowers, mortgage activity, and trends in loan amounts (Pettit and Droesch 2008). The availability of these indicators spurred a substantial body of research on the mortgage markets and empowered activists to reveal racial disparities in lending. When combined with a list of lenders specializing in subprime lending, this dataset played a critical role in documenting the spatial and racial patterns of subprime loans during the late 1990s and early 2000s (Immergluck and Wiles 1999; Treskon, Kornil and Silver 2003). Clearer definitions implemented in 2004 improved the ability to identify these loans from all lenders. Beginning in 2009, HMDA data were distributed online.

Another example of a new data series that is useful for neighborhood and community information is the US Census Bureau's Longitudinal Employer–Household Dynamics, which was launched in 2002. This data series uses sophisticated statistical and computing techniques to combine federal and state administrative data on employers and employees with other Census Bureau data, maintaining confidentiality protections for individuals and firms. Under the Longitudinal Employer–Household Dynamics, the Local Employment Dynamics Program is a voluntary partnership between state labor market information agencies and the US Census Bureau to develop new information about local labor market conditions. By relying on existing administrative data, the Census Bureau produces the new dataset at low cost and with no added respondent burden (Dowell n.d.). The partnership began with 10 states and by 2011 had expanded to all 50 states. In 2005, the Local Employment Dynamics Program debuted OnTheMap, a data series showing where workers are employed and where they live. The development of OnTheMap demonstrated remarkable advances in politics, methodology, and dissemination of data. The data series shows the payoffs from the cooperation between state government agencies and the federal government to create a new data resource. OnTheMap was a methodological innovation because it was the first synthetic data product released by the Census Bureau. The Bureau uses state unemployment insurance records, individual wage records, and ES-202 records on employers to create a dataset with similar statistical properties to the original dataset but without compromising confidentiality (Lane et al. 2001). The data series is published at the census block level and consists of three components: (1) counts of jobs by industry, earnings, and worker characteristics such as race, age, sex, and educational attainment; (2) counts of workers by their residential location for the same categories; and (3) a commuting series that reports worker counts for pairs of residential blocks and employment blocks. The Bureau's dissemination efforts are also extraordinary. Users can access OnTheMap data through a powerful online mapping and reporting tool, and a listserv and annual conference have cultivated an active user community for peer learning and support. (See the Wascalus and Matson essay at the end of chapter 2 for a description of Minnesota-3D, a local portal that allows users to visualize patterns from Local Employment Dynamics data.)

In another illustration, the US Department of Housing and Urban Development (HUD) recognized the need for more current, fine-grained data about property vacancy as the foreclosure crisis began to unfold in early 2007. The US Postal Service maintains records on the status of mail delivery for every address in the country, and these records contain information on whether the house is vacant. However, the Postal Service sells the records to private firms to repackage, primarily for marketing purposes, and the cost of purchasing these records is too high for most public interest groups. HUD successfully negotiated with the Postal Service to publish census tract–level counts of address by mail delivery status at no cost to the public. (In a subsequent round of negotiations ending in 2012, HUD restricted access to governments and nonprofit organizations.) The data played an important role in identifying tracts eligible for the later rounds of the Neighborhood Stabilization

Program (US Department of Housing and Urban Development 2010b, n.d.b). These data represent an important precedent in transforming commercially valuable data to a less-detailed form in the public interest. These data are also significant for their timeliness and frequency. HUD publishes the data quarterly and with less than two months' lag, vastly improving their value to policymakers and housing practitioners. The data still exhibit some of the downsides of proprietary data. For example, there is no documentation on changes in methodology that cause swings in indicator values, nor is there a way for researchers to submit inquiries to the producers of the data.[1]

In December 2010, the United States entered a new data era when the US Census Bureau released the first set of five-year American Community Survey (ACS) data. The Census Bureau designed the ACS to replace the former long form of the Decennial Census and increase the frequency of updated data. ACS provides annual releases of data so that users don't have to wait 10 years for a Decennial Census update.[2] The Bureau releases one-year data for geographic areas with a population of 65,000 and higher; three-year data for areas with a population of 20,000 and higher; and five-year data for all areas, including block groups and census tracts. Because the ACS is now administered every month, the Bureau can maintain a consistent level of operations compared with the drastic ramp-up previously needed for the long form. This change means that the Bureau can maintain a set of long-term, trained staff to execute the survey, resulting in improved data collection practices.

Users of all technical abilities are struggling to learn how to responsibly analyze, interpret, and communicate indicators based on the five-year ACS data. The obvious change is that the five-year ACS data are a period estimate, reflecting conditions over the entire time period, in contrast to the point-in-time estimate in the long-form. This difference would be difficult to communicate to a lay audience at any time, but the problem was compounded by the timing of the inaugural five-year estimates. The years 2005 to 2009 straddled the end of a boom period and the start of the housing crash and Great Recession, so indicators on economic conditions and housing values did not match with the general understanding of conditions at the time of the data release.

Users faced another learning curve in the need to consider margins of error. Margins of error existed for all long-form Decennial Censuses, but the Census Bureau did not publish them or emphasize their use. The lower sampling rate in the ACS results in higher margins of error than those obtained with the higher sampling rate used in the Decennial Census long form. For the first ACS release of census tract–level data based on surveys collected from 2005 to 2009, the sampled households represented only 11 percent of all households, compared with about 17 percent in the 2000 Decennial Census (Hefter 2010). This difference resulted in coefficients of variation for ACS estimates that were roughly 1.5 to 2 times larger than their census sample counterparts (Beaghen et al. 2012). Starting in 2011, the Census Bureau increased the annual target sample from 3 million addresses to 3.54 million, which will improve the sample coverage somewhat over time.

The Census Bureau now prominently publishes the upper and lower limits alongside the estimate value. Even so, the estimates are generally presented in news articles or community analysis without regard to margin of errors. For analysis purposes, the larger margins of error make it more difficult to confidently differentiate among conditions in different geographies in a given year, or across years for a given geography. More troubling is that the reliability of the estimates is not constant across different types of neighborhoods. The coefficients of variation are higher for census tracts that have lower median household income and greater shares of nonwhite population, creating particular challenges for the study of vulnerable neighborhoods (Spielman et al. 2014).

Recognizing the difficulty users would have transitioning to the ACS, the Census Bureau produced the Compass products, a set of educational materials produced by Bureau staff and contractors. These materials include several handbooks for different audiences (e.g., researchers, journalists), presentations on a variety of topics, and an e-tutorial (US Census Bureau n.d.). In addition, the Census Bureau funded the Population Reference Bureau in early 2013 to launch an ACS user's group, which will provide training at academic conferences and establish an online community platform where users can exchange questions and resources. In addition, several Census Information Centers and NNIP partners have developed training for their nontechnical audiences on using and understanding the ACS.

## State and Local Government Data

Chapter 2 provides an introduction to administrative data generated by state and local government operations and record keeping, as well as the increased internal use of the systems to improve efficiency and accountability. Coulton (2008) provides a comprehensive overview of administrative data sources. She documents issues to consider when repurposing administrative data, including the lack of metadata and the need to protect the confidentiality of individuals' data. The report provides sample indicators and likely source agencies for data files in eight categories: economy, education, health, social services, safety and security, community resources and participation, housing, and environment. In addition, NNIP has published guidebooks over the years that give advice and examples on specific data sources.[3]

Multiagency program service data, consisting of client-level observations of publicly funded services, represent another type of administrative data. These data may be entered by a city agency providing a service directly or by a collection of nonprofits that is contracted by local governments to provide services, such as mental health or child welfare services. A prominent example of these data is the Homeless Management Information System, required by HUD to be maintained by local Continuum of Care lead organizations (US Department of Housing and Urban Development n.d.a). The Homeless Management Information System results from a 2001 Congressional mandate that HUD work with jurisdictions to gather more detailed homeless data by 2004. Although these data are less commonly used to create neighborhood indicators, they often contain addresses and could provide insight into spatial patterns of service provision or supportive housing.

Many issue areas have seen impressive progress over the last decade in increasing access and reuse of administrative data for neighborhood indicators. NNIP partners, for example, use data from many local and state government sources and turn them into useful indicators. (See table 3.2 for an NNIP partners' data inventory.) An illustration of how state and local data have been used by multiple partners comes from the Reentry Mapping Network, a collaborative effort by the Urban Institute and community-based organizations in 15 cities. At the outset of this project, no NNIP partner organization had access to administrative records of data from prisons or jails. However, community organizations recognized that having individuals, particularly young men of color, cycling through incarceration and reentry hampered their improvement efforts in low-income neighborhoods. In response to these concerns, the Reentry Mapping Network was designed to instigate community change through the mapping and analysis of neighborhood-level data on prisoner reentry. Reentry Mapping Network partners collected and analyzed

**Table 3.2.** NNIP Partner Data Holdings, 2013

| | Total | Address/ parcel | School | Other small area identifier |
|---|---|---|---|---|
| **Births and deaths** | | | | |
| Births total | 27 | 14 | 0 | 13 |
| Births by prenatal care | 26 | 12 | 0 | 14 |
| Births by birth weight | 26 | 13 | 0 | 13 |
| Deaths by cause | 20 | 10 | 0 | 10 |
| **Education** | | | | |
| Student enrollment | 34 | 13 | 16 | 5 |
| Student proficiency | 35 | 11 | 19 | 5 |
| Student absences | 34 | 14 | 15 | 5 |
| Free/reduced price lunch | 35 | 12 | 18 | 5 |
| Special education | 26 | 10 | 10 | 6 |
| Kindergarten readiness assessment | 14 | 5 | 4 | 5 |
| Head Start enrollment | 14 | 6 | 6 | 2 |
| Other pre-school enroll. (by type) | 16 | 5 | 7 | 4 |
| Licensed child care | 24 | 21 | 3 | 0 |
| **Health** | | | | |
| Immunization | 9 | 1 | 1 | 7 |
| Child blood-lead level | 13 | 4 | 0 | 9 |
| Hospital admissions by cause | 14 | 2 | 0 | 12 |
| Asthma hospitalizations | 13 | 3 | 0 | 10 |
| Emergency department visits | 8 | 2 | 0 | 6 |
| Ambulatory care | 3 | 0 | 0 | 3 |
| Injury surveillance data | 3 | 0 | 0 | 3 |
| Communicable diseases | 6 | 1 | 0 | 5 |
| Sexually transmitted diseases | 7 | 2 | 0 | 5 |
| **Public assistance** | | | | |
| TANF | 14 | 4 | 0 | 10 |
| Food stamps | 11 | 4 | 0 | 7 |
| Medicaid | 11 | 6 | 0 | 5 |
| S-Chip | 5 | 3 | 0 | 2 |
| WIC | 8 | 3 | 0 | 5 |
| Subsidized child care | 8 | 4 | 3 | 1 |

*(continued)*

**Table 3.2.** *(Continued)*

| | Total | Address/ parcel | School | Other small area identifier |
|---|---|---|---|---|
| **Housing assistance** | | | | |
| Public housing units | 19 | 17 | 0 | 2 |
| Housing choice vouchers | 15 | 11 | 0 | 4 |
| Other subsidized housing | 15 | 11 | 0 | 4 |
| **Crime** | | | | |
| Reported crime (Part I) | 25 | 21 | 0 | 4 |
| Reported crime (Part II) | 22 | 18 | 0 | 4 |
| Arrests | 16 | 12 | 0 | 4 |
| Arrests (juvenile) | 14 | 8 | 0 | 6 |
| Emergency (911) calls | 9 | 7 | 0 | 2 |
| **Prisoner reentry** | | | | |
| Ex-offenders returning from prison | 14 | 9 | 0 | 5 |
| Ex-offenders returning from jail | 9 | 6 | 0 | 3 |
| Persons on probation/parole | 10 | 8 | 0 | 2 |
| **Business/economy** | | | | |
| Business inventory (ES-202) | 15 | 5 | 0 | 10 |
| UI claimant file | 4 | 1 | 0 | 3 |
| Business inventory (other) | 17 | 13 | 0 | 4 |
| Business licenses | 7 | 7 | 0 | 0 |
| Liquor licenses/stores | 17 | 17 | 0 | 0 |
| **Property transactions/characteristics** | | | | |
| Property characteristics | 27 | 27 | 0 | 0 |
| Property sales (volumes, prices) | 26 | 23 | 0 | 3 |
| Property tax assessments | 32 | 31 | 0 | 1 |
| Tax delinquencies | 19 | 19 | 0 | 0 |
| Evictions | 4 | 3 | 0 | 1 |
| Vacant parcels | 25 | 20 | 0 | 5 |
| Foreclosures | 33 | 24 | 0 | 9 |
| Building permits | 22 | 20 | 0 | 2 |
| Demolitions | 21 | 19 | 0 | 2 |
| Housing code violations | 20 | 18 | 0 | 2 |
| Lead paint abatements | 5 | 2 | 0 | 3 |
| Water usage | 5 | 4 | 0 | 1 |

**Table 3.2.** *(Continued)*

| | Total | Address/ parcel | School | Other small area identifier |
|---|---|---|---|---|
| Water shutoffs | 7 | 6 | 0 | 1 |
| Electric shutoffs | 3 | 2 | 0 | 1 |
| **Child welfare** | | | | |
| Foster care | 6 | 4 | 0 | 2 |
| Child abuse/neglect | 10 | 4 | 0 | 6 |
| **Other** | | | | |
| Voting records | 17 | 11 | 0 | 6 |
| Community referral calls | 10 | 5 | 0 | 5 |

*Source:* NNIP Data Inventory Results. Published January 2014.

local data related to incarceration, reentry, and community well-being and worked with multiple sectors to craft strategies for addressing prisoner reentry in their communities. The project culminated with a guidebook that framed the issues, gave detailed information on accessing and processing the data sources, and used the experiences of the Reentry Mapping Network partners to illustrate lessons about how to develop successful reentry strategies (LaVigne, Cowan, and Brazzell 2006).[4]

Another important development has been the use of indicators based on property data for program planning and policymaking. Although the original NNIP partners had property-related data in the mid-1990s, the field had advanced considerably by 2012, with more frequent analysis of property characteristics and status from a wider range of data sources. These improvements were enabled by better systems at the local government level. By 2005, almost three-quarters of America's top 100 cities had digitized parcel-based information systems, and they made a considerable amount of their data available to the public via the web (Chandler et al. 2006). A cross-site project conducted by NNIP (Kingsley and Pettit 2008) shared examples of using property data for decisions about government actions and community development investments in Atlanta, Georgia; Baltimore, Maryland; Indianapolis, Indiana; Milwaukee, Wisconsin; Providence, Rhode Island; and Washington, DC. Parcel data with land and building characteristics from assessors' offices served as the backbone for most of the local analysis, but NNIP partners used many other sources of data on properties, including foreclosure filings and sales, subsidized housing records, housing code enforcement records, and program data from community development corporations and the agency managing the Community Development Block Grant.[5] The new property-specific applications of parcel data required more frequent updates of the data. In a growing number of cities, these types of data are now updated at least quarterly to support timely local action toward neighborhood stabilization and community development goals.

While analysis combining parcel-level data sources grew more sophisticated, other institutions in and outside of government were designing complex systems to merge person-level data from multiple agencies. These integrated data systems (IDSs) link individual-level records from multiple administrative agencies on an ongoing basis while maintaining confidentiality requirements (US Government Accountability Office 2013). A survey by a coalition of national organizations identified 30 operational IDSs across the country in early 2013: 5 at the state level and 25 at the county or city level (Urban Institute 2013). The systems are created and maintained under a variety of institutional arrangements. For example, Chapin Hall at the University of Chicago manages the Illinois Integrated Database on Child and Family Programs, and the Providence Plan developed the Rhode Island DataHub (Providence Plan 2013; Goerge 2011; Kitzmiller 2013). The data from participating agencies vary, but they include data on health, employment, birth certificates, education, justice system involvement, and human services (such as child welfare, income supports, and child care subsidies). IDSs are used to assist in compliance and program supervision, to improve case management for individuals, or to inform policymaking and program evaluation.

Actionable Intelligence for Social Policy, a network based at the University of Pennsylvania, was created to support and advance IDSs across the country. This network documents the development and use of existing systems and is developing tools and guides to help communities create new ones. Many other national organizations, including the National League of Cities and the Data Quality Campaign, have recognized the importance of these systems for their constituencies and are also documenting practices and conducting trainings on aspects of IDSs. Although the records that are incorporated into an IDS may contain residential addresses, few examples exist of using these systems for neighborhood-level analysis or of considering neighborhood-level conditions in individual-level analysis. To incubate more examples of using an IDS with a neighborhood lens to inform local policy, in 2013 the Annie E. Casey Foundation funded several local organizations in an NNIP cross-site project to support such projects (Hendey, Coulton, and Kingsley 2013). The growth of IDSs over the past few years suggests that these systems will continue to expand, both in number and variety of uses.

### *Proprietary Data*

The availability of proprietary small-area data increased concurrent with the rise in federal data releases. Commercial firms, universities, or nonprofits own the rights to these data and sell or share the data with restrictions on access and/or redistribution. These products are generally created to serve the analytic needs of other commercial enterprises, such as evaluating a business location or identifying a likely area to market goods and services. Proprietary data expand the information available for neighborhood indicators, but they can present difficulties for community groups and even researchers to access and use. Cost and license restrictions against redistribution are the two obvious barriers, but not the only ones. Firms selling products have few incentives to reveal the detailed methodology they use for computation and estimation. Sales people may exaggerate the quality or coverage of their proprietary datasets, and users do not often have the budgets to purchase and evaluate comparable products. Despite these limitations, commercial data products have proved useful in some instances for understanding neighborhood change and informing local policy and programs. A few proprietary data products on people, businesses, and property that are relevant to community planning and public policy are described below.

Companies often repackage or add value to public federal data and publish them in usable forms. For example, the Urban Institute partnered with GeoLytics, Inc. in the early 2000s to create the Neighborhood Change Database. Funded in part by the Rockefeller Foundation, the Neighborhood Change Database updated the Urban Institute's Underclass Database from the 1980s. The Neighborhood Change Database has several features that make it easier for researchers to study neighborhoods over time. First, using a publicly documented methodology, it weights the data from the 1970 to 1990 Decennial Censuses to represent census tract boundaries as of 2000. Second, it relieves users from having to look up which tables and cells to add together by creating a set of standard variables, such as poverty rate, that are equivalent across the four surveys. It also offers one of the few sources for tract-level data from 1970 and 1980, which are no longer obtainable from the Census Bureau. An updated product including the 2010 Decennial Census ACS five-year averages for 2006 to 2010 was released in 2014. Researchers from the Urban Institute and other institutions use the Neighborhood Change Database to study many aspects of neighborhood change.

Many companies publish annual estimates of basic demographic, social, and economic conditions for census tracts or Zip Codes; these estimates were initially intended to serve decisionmakers in the years between the long form of the Decennial Census. In the new environment created by the ACS, these private data still play a role in generating annual estimates from the five-year ACS data, supplemented by other national or local datasets. Nielsen Claritas is the most well-known firm in this market, but Esri and many other firms are active. Another source of data that has been aggregated for neighborhood indicators is household credit scores from sources such as Equifax or Experian. Researchers have summed person-level records to create neighborhood indicators. In one example, the share of residents with high credit risk was tested as a predictor of foreclosures (Grover et al. 2008).

Business listings are another common category of commercial data products. These data, which include characteristics such as location, industry classifications, and business size, are derived from a combination of sources such as credit applications, telephone directories, state business license lists, and Internet research. Business information from firms such as Dun and Bradstreet or InfoUSA has been used to create neighborhood indicators of residential services (both positive and detrimental). In one example, described in more detail in the section on neighborhood scores, researchers from The Reinvestment Fund (TRF) used data from Trade Dimensions to identify supermarket locations and assess neighborhood access to supermarkets (Califano et al. 2011). In another example, a Brookings Institution analysis explored and compared the prevalence of high-cost nonbank basic financial services in low-income versus high-income neighborhoods (Fellowes and Mabanta 2008).

Other prominent examples include property-related data compiled from a variety of local government and commercial records (Pettit et al. 2014b). The US Postal Service vacancy data, mentioned earlier because they are released by HUD as aggregate counts, are also available quarterly for a fee at the individual address level. Companies like Experian and

First American have assembled and resold property characteristics and sales data from local governments since the 1990s. These data files have been used for many research studies on the home price impacts of various interventions (Galster et al. 1999; Galster and Tatian 2009).

The foreclosure crisis in the late 2000s revealed the weakness in accessible fine-grained data for analysis and decisionmaking. Firms such as RealtyTrac began to compile local government agency and court records, both in electronic form and through digitizing PDFs or even paper records. Their primary customer base was investors seeking to buy foreclosed properties. Repeated purchases of these proprietary datasets were not feasible for local governments and nonprofits needing to make property decisions or for many researchers interested in documenting the cause and effects of foreclosures. A related type of data tracks mortgages instead of properties. LPS, formerly McDash Analytics, founded its business model on collecting mortgage performance records from servicers in order to produce industrywide analytics and benchmarks. These data became sought after to understand patterns of delinquency and foreclosure. In general, loan-level data from this source for the country cost hundreds of thousands of dollars, and Zip Code–level data prices were tens of thousands. The Federal Reserve and a limited number of researchers had the means to purchase these datasets and conduct nationwide research (Chan et al. 2011; Mayer et al. 2011a, 2011b), but the high price sharply curtailed its usefulness for local policy and programs.

## Primary Data Sources

Administrative data are shaped by the purpose of their collection and include the data that are necessary for organizations to operate programs or regulate activities. Often, communities need information about neighborhoods that is not captured by government records and turn to collecting primary data to suit their specific purposes. Primary data can have many advantages, including potentially being more timely, capturing intangible information, and describing a narrow target area. Primary data also have drawbacks. Whether through volunteer effort or direct costs, collecting primary data can be costly in time and money. Ensuring data consistency and quality requires training of the data collectors, and updating the data requires another round of data collection. Finally, there may not be comparable indicators for other neighborhoods or in other cities against which to benchmark the results. This section describes two primary data collection methods often used to create neighborhood indicators: observational surveys and household surveys.

### Observational Surveys

Observational surveys involve individuals visually identifying a set of neighborhood or property conditions and completing a protocol to record the data. The rigor of data collection varies widely, but the ideal process includes clear definitions of the characteristics being captured, training of the data collectors, and data-quality checks. These surveys have been used to capture deficits in a neighborhood, such as graffiti or trash, as well as community assets, such as parks and fresh food outlets.

Observational surveys have often been used to compensate for the lack of sound information on vacancy and property conditions, as even the most advanced city agencies have trouble collecting these data through administrative systems or self-reporting. In Detroit, Michigan, the Detroit Residential Parcel Survey offers one example of how surveys successfully filled this critical data gap. Although the extreme problem of vacant and abandoned buildings in Detroit was undisputed, no citywide information existed before 2009 to develop scenarios and make informed decisions about different areas of the city. In response to this need, the survey project was conceived and implemented by the Data Collaborative, a partnership between the Detroit Office of Foreclosure Prevention and Response, Community Legal Resources, and Data Driven Detroit (the local NNIP partner). In two months in 2009, surveyors recorded the conditions of every residential property with one to four housing units in the city of Detroit, totaling about 350,000 structures. The surveyors recorded property type, property condition, fire damage, lot vacancy, and type of structures. The collaborative published the results in static and interactive maps accessible to the government, community groups, and the public. Data Driven Detroit took advantage of more sophisticated technology to update the parcel survey in 2014, and their plans included ways to keep the information continuously updated (Detroit Blight Removal Task Force 2014).[6]

In addition to obtaining information about properties, observational surveys are often used to measure walkability and the quality of the built environment. The Irvine Minnesota Inventory is a well-tested audit tool for capturing features that promote physical activity, particularly walking (Day 2006). Developed and refined by the University of California, Irvine, and the University of Minnesota, the questions cover four domains: accessibility, pleasurability, perceived safety from traffic, and perceived safety from crime. Examples of the questions' topics include sidewalk features, traffic signs, and land use. The protocols and training presentations are available online.[7]

### *Household Surveys*

Often the information needed to understand community issues cannot be captured by visual surveys alone. In addition to demographic, economic, and social characteristics, household surveys allow analysts to capture residents' motivations and priorities, their opinions of the neighborhood, and their views on the drivers of community conditions. Examples exist from one-time surveys for a single neighborhood to long-term metropolitanwide surveys. Like observational surveys, neighborhood surveys have a wide range of purposes and statistical rigor.[8]

Two surveys are illustrative of exemplary survey design and have introduced innovative measures of neighborhood characteristics. The Project on Human Development in Chicago Neighborhoods researched the various factors that affect child and adolescent development, and the team conducted a community survey as part of the study (Earls et al. 1995).[9] The survey measured the social and physical conditions in a probability sample of neighborhoods in Chicago over 1994 and 1995. The researchers' concepts of informal and formal social control and social cohesion have been particularly influential on the field, and the questions related to those ideas have been included in many subsequent surveys.

The Los Angeles Family and Neighborhood Study sought to determine the characteristics of supportive neighborhoods in Los Angeles County, California.[10] In particular, it focused on various factors in children's development, welfare reform effects on neighborhoods, and residential mobility and neighborhood change (Rand Corporation n.d.). Its community survey was first conducted in 2000 and 2001, with the latest wave in 2006 and 2008. The survey, conducted in a representative sample of neighborhoods, covered many topics: household composition, social and economic indicators, health status, participation in public and private assistance programs, mobility, and social ties and social support. Unlike many surveys, the survey followed both neighborhoods and families over time. This meant that families who were interviewed in the first round were interviewed in the second, even if they had left the neighborhood. The survey also included a sample of newcomers. Another survey with this feature was fielded to support the Annie E. Casey Foundation's Making Connections initiative in low-income neighborhoods in 10 cities. The initiative and the contributions of the survey are described in depth in chapter 5.

The Boston Neighborhood Survey, a large-scale survey conducted from 2006 to 2010 that produced rich neighborhood indicators, was a biennial telephone survey of adults that had a sufficient sample size to produce indicators for Boston neighborhoods. The survey asked respondents about topics such as community norms and neighborhood resources, respondents' sense of community well-being and perceptions of community safety, demographic characteristics, and the well-being of neighborhood youth.

## Notable Data Trends

The previous section documented the progress made in the quantity of secondary and primary data available for local analysis since the founding of NNIP in the mid-1990s. In addition, several recent trends are expanding the variety and accessibility of data. In this section, we discuss four of these developments that have particular relevance for local stakeholders. *Neighborhood scores and metrics* demonstrate how analytics can produce concise and innovative measures for key quality of life issues. Next, *big data*, from sources such as social media and physical sensors, offer new types of information that have potential to generate insights about neighborhood conditions and resident interests. Third, *open government data* are increasing the amount of easily accessible data that can help improve policy and practice, as well as support government transparency efforts. Finally, the efforts to develop *data standards* for local administrative data could lead to more comparable data and consistent neighborhood measures over time. Viewed together, the four trends indicate that the next decade will see exciting advances in neighborhood data for community action.

### *Neighborhood Scores and Metrics*

Private firms, nonprofits, and government agencies are increasingly using small-area government and commercial data to create nationally available metrics, which may be derived from a single measure or from multiple ones. Chapter 7 discusses the methodological aspects of composite indices. This chapter reviews examples of how organizations designed new indicators in an easy-to-access format on topics of interest to a broad group of organizations and the media. By publishing the actual data (not only a narrative analysis), these organizations successfully framed their issue area and drew the attention of advocates, policymakers, and individual households.

The best known example is Walk Score, an indicator ranging from 0 to 100 that measures walkability based on walking routes to destinations such as grocery stores, schools, parks, restaurants, and retail. The score's visibility has spread as realtor sites like Zillow incorporate it into home listings, government agencies require its use in grant applications, and cities use it to plan investments (US Department of Housing and Urban Development 2010a; Walk Score n.d.). The company has also developed Transit Score and Bike Score and continues to create new analytics for rating neighborhoods. The methodology is proprietary, but there is a process for researchers to access Walk Score data that has enabled research about its validity (Duncan et al. 2011; Carr, Dunsiger, Marcus 2011).

Another example is the influential Housing + Transportation Affordability Index introduced in 2006 by the Center for Neighborhood Technology, mentioned below as a pioneer in Neighborhood Early Warning Systems (Center for Transit-Oriented Development and Center for Neighborhood Technology 2006). The Housing + Transportation Affordability Index reflects both housing and transportation costs and sets the acceptable limits for these costs at 45 percent of household income. The index has contributed to an increasing awareness of a more comprehensive approach to affordability and a better picture of economic hardship for households (Urban Land Institute 2006; Hickey et al. 2012). The index also is consistent with views of advocates countering policies that promote sprawling development, which may offer lower-cost housing if transportation is not considered. The index is calculated for block groups in over 330 metropolitan areas in the United States based on housing cost data from the ACS and a model capturing three components of transportation behavior: auto ownership, auto use, and transit. Using the Housing + Transportation Index as a model, HUD's Office of Sustainable Housing and Communities, in partnership with the Department of Transportation, released its own Location Affordability Index with visualization tools in 2013 (US Department of Housing and Urban Development 2013a).

The final examples of developing a framework to combine individual indicators are organizations that seek to measure the extent to which neighborhoods have access to full-service grocery stores and, in particular, to identify "food deserts." In 2011, TRF, a community development financial institution, published an analysis on the Limited Supermarket Access Area Score (Califano et al. 2011). The authors defined a limited supermarket access area as one in which residents must travel significantly farther to reach a supermarket than the comparatively acceptable distance traveled by residents in well-served areas.[11] TRF defines *comparatively acceptable* as the distance that residents of well-served areas (block groups with incomes greater than 120% of the area's median income) travel to the nearest supermarket. The data sources included Trade Dimensions for supermarket locations; the Decennial Census for population, households, and residential land area; ACS data for household income; and the Bureau of Labor Statistics Consumer Expenditure Survey for demand for food at home. TRF also published block group data for the nation on their PolicyMap data portal (described below). In 2013 the US Department of Agriculture introduced complementary data and the Food Access Research Atlas, an online mapping tool. In their definition, a census tract is considered to have low access if a significant number or share of individuals in the tract is far from a supermarket.

### *Big Data*

The administrative and primary data sources listed above are similar to those available to NNIP partners in the mid 1990s. The technological advances described in the next section create the potential for new types of data sources for neighborhood indicators. Many of the new sources loosely fall under the umbrella of "big data," which refers to data that have levels of volume, velocity, and/or variety that traditional computational techniques cannot handle. The commercial sector already leverages big data for marketing, and Fleming's essay at the end of chapter 2 sets out aspirational goals for wider use of big data by government. Universities have developed specialized centers to advance analytic techniques using big data to the benefit of the public sector and wider community. These centers include the Center for Urban Science and Progress at New York University, the Urban Center for Computation and Data at the University of Chicago, and the Event and Pattern Detection Laboratory at Carnegie Mellon University.

Big data come in many shapes. Data from social media, such as Twitter feeds or mapping FourSquare check-ins, are the most visible. The Livehoods Project, a research project from the School of Computer Science at Carnegie Mellon University, is combining social media data and machine-learning techniques to develop new methodologies to portray social patterns across cities. Other data, like the geographic features in Open Street Map, may be crowdsourced, that is created by many distributed contributors (MapBox 2013).

Another type of big data is imagery; infrared pictures can map the surface temperature across the city (Environmental Protection Agency n.d.).

Mobile devices are now a common tool for conducting observational surveys and collecting other primary data. SeeClickFix, a private firm, is one national example. This mobile application (app) allows anyone to report nonemergency issues (e.g., potholes, broken streetlights) to their local governments; local governments have also developed their own apps that can upload photographs and record geographic coordinates to submit with the request. Cell phones and mobile tracking devices can also generate data themselves, charting preferred routes for a bicyclist or the volume of communications throughout the day. Physical sensors can provide another rich source of data, measuring such diverse items as air quality and automobile and pedestrian traffic. The potential for better understanding neighborhood patterns through data from mobile phones and other sensors will increase as these devices are connected to the Internet, forming the Internet of Things (IoT).

However, the applications of big and unstructured data for public policy in general, and neighborhood indicators in particular, are more promise than practice at this point. The barriers include the obvious lack of advanced technical knowledge, but also the need to learn more about how the new indicators could inform action on community issues. These new data also introduce the risk of distorting the real picture if not interpreted correctly, because social media participation varies by demographic and economic groups. They also raise concerns about privacy protections, especially for data mining being conducted outside of universities without structured review processes. Even with these hurdles, the insights to be gained from new sources of data provide enough incentive for researchers and practitioners to learn how to collect, organize, and interpret them, and their use will undoubtedly gain momentum in the coming years.

## *Open Government Data*

The open government data movement is founded on the premise that technology can improve the transparency of processes and information so that citizens can hold governments accountable. Advocates view government data as a public good that should be available to the taxpayers that funded their creation. Open data can also encourage citizen engagement in government decisionmaking. Another argument for open data is the added value that government data can bring to the private sector and society in general.[12] Sanders's essay at the end of this chapter describes the advancement of this idea by the Obama administration and local governments, noting it is just one component of the larger agenda for delivering relevant information to communities. Large cities like Chicago; San Francisco; Philadelphia, Pennsylvania; and New York have taken the lead in open data, but the efforts of groups like the Sunlight Foundation and the Open Knowledge Foundation have promoted the ideas more broadly. As of summer 2013, data.gov listed open data projects from 40 states and 41 counties and cities in the United States.

This movement will continue to increase the availability of nonconfidential data relevant to neighborhood indicators, notably reported crime, 311 service requests, and property-related data. Right now, the volume of raw data means that data analysts or software developers are the primary users of these systems. Cities rarely publish extensive documentation or even basic metadata for the files. The tasks of cleaning the data and crafting appropriate indicators still remain for local data intermediaries and other analysts, but readily available access will save users time and promote cross-city analysis (Pettit et al. 2014a).

## *Data Standards*

The development of standards (i.e., defined structures for different types of data) is another growing trend related to government data that could facilitate the availability and use of neighborhood indicators. In mid-2009, the District of Columbia government and entrepreneurs from SeeClickFix developed specifications for data in 311 systems, centralized call centers, and online systems maintained by local governments for residents to submit service requests for problems such as potholes, bulk garbage pickups, or broken streetlights.[13] With the participation of the organization Open Plans and other pioneer cities, the specification has been generalized and improved since its initial conception. In 2013, the Open311 website listed about three dozen cities that had adopted the standard. Fleming's essay describes the role of 311 data overall in improving local government operations.

The General Transit Feed Specification offers another impressive example of data standards; more than 200 local governments are using its published data on transit stops, routes, trips, and other schedules.[14] The adoption of this standard most prominently powers the public transit directions on Google Maps, but it is open source and can be used by any application. One more recent effort is HouseFacts,[15] a proposed standard for health and code violations of residential properties. The standard was developed by a coalition consisting initially of a private-sector firm that offers code enforcement management software, the City of San Francisco, and Code for America. The short-term hope is to have unified property profiles for prospective homebuyers or renters to evaluate the quality of homes they are considering. Longer term, the standard has the potential to deter violators through more public scrutiny and help code enforcement agencies better manage internal systems.

How do these standards relate to neighborhood indicators? The spread of standardized local data facilitates application development for individual needs, but it also opens up opportunities for data analysts, researchers, and practitioners. As more cities come on board, the likelihood of consistent definitions of data fields over time increases. This comparability could enable the development of indicators for larger geographic areas and easier benchmarking of neighborhood conditions across cities. For example, one could aggregate data from multiple jurisdictions up to a metropolitan-level indicator for the share of households within one-half mile of a bus stop or compare the rate of service calls from low-income neighborhoods in one city to a parallel figure in comparable cities. Like open data, standards also lower barriers for academic research that includes multiple cities.

## Technology Related to Neighborhood Indicators

Technological advances since the mid-1990s have changed many aspects of creating and using neighborhood indicators. More powerful and affordable hardware and software and the growth of the Internet have expanded neighborhood indicators in scale and scope. Sanders's essay describes the powerful tools and infrastructure available today to collect, store, and share data, including application programming interfaces (APIs), mashups, geolocation, and mobile technology. The essay also includes recommendations about the additional investments in technology and skills that are needed for communities to benefit from the advances. The new powerful equipment and applications have transformed both the backend production and processing of neighborhood indicators and the more visible data delivery and visualization.

## Production and Processing

Production refers to the creation of the initial data that form the basis of neighborhood indicators. Today's technology has increased the speed of traditional data collection and expanded the types of data collected. With automation now commonplace, delays due to transfer of paper records to electronic systems are eliminated. For example, social service workers stationed offsite at food pantries enter eligibility criteria into the Supplemental Nutrition Assistance Program information management systems, and residents submit city service requests online through 311 systems. The technology also facilitates the collection of new types of data. Collection of data by mobile devices automates the collection of Global Positioning System (GPS) coordinates and may include the option for an accompanying photo.

The more efficient production of data facilitates easier access to the data for both government agency staff and external users. In a few cases, NNIP partners receive automated daily feeds of administrative data, a far cry from the cumbersome reels of data tapes used in the early years. The technology available today would not have been imagined by the founding NNIP partners. The new specifications sharply reduce the production time needed to convert raw data into neighborhood indicators.

Although many types of software have improved in the past two decades, desktop geographic information systems (GIS) deserve a special mention. In 1990, the US Census Bureau released the first Topologically Integrated Geographic Encoding System (TIGER) files. This revolutionary database contained information to represent the location of roads, legal boundaries, and other map features for the entire United States (Klosterman and Lew 1992). With TIGER and GIS software, analysts could assign map coordinates to addresses from administrative records, assign identifiers for census tract or neighborhood boundaries, and aggregate the data to create neighborhood indicators. The pairing of this data and technology certainly was a prerequisite for the NNIP model, and it also affected local governments, as related in Fleming's essay in chapter 2. Esri currently dominates the desktop GIS market in the public sector with its ArcMap software, but open source mapping software and online services are growing in popularity. For now, the wide adoption of Esri's products provide a default format for file sharing and student and staff training.

## Data Portals: Delivery, Visualization, and Analysis

In addition to facilitating the collection and processing of data, technology has resulted in new tools to share, visualize, and analyze data. Since 2010, the number of data and indicators portals has exploded, and continuous advances in available technology have led to a constantly shifting environment. The following discussion is organized by the type of organization that stewards the data portals, with an emphasis on the growth in nonprofit systems.

### *Nonprofit Portals*

Local data intermediaries pioneered the field of online community information in the late 1990s with websites that organized data and indicators to support community action. One set of early systems shared a common purpose: to assemble data on individual properties in one place in order to predict and combat financial disinvestment and building abandonment. Snow et al. (2003) documents the origin and contents of these "early warning systems" in Chicago, Los Angeles, Minneapolis, and Philadelphia.[16] The latter three emerged in the late 1990s, but the Center for Neighborhood Technology led the way more than a decade earlier by unveiling Chicago's Neighborhood Early Warning System in 1984 using shared floppy disks.[17] These systems combined community-collected data with administrative data on land use, code violations, water arrears, and property tax delinquencies. They enabled community development corporation staff, city officials, and residents to access an unprecedented amount of integrated information to track problem properties through user interfaces tailored for their needs.

Early warning systems largely focused on detailed parcel information, but other early data tools centered on aggregated local data.

Case Western University in Cleveland, Ohio, a founding NNIP partner, launched the Community and Neighborhood Data for Organizing (CANDO) in 1992 as a system that users dialed into by modem. The system contained social and economic neighborhood indicators based on the 1990 Decennial Census and local administrative sources. The Urban Institute 1996 analysis of the six original NNIP partners mentions three of the partners using the web to distribute data: the Cleveland site, the Piton Foundation's Neighborhood Facts in Denver, Colorado, and The Providence Plan neighborhood profiles.

During the decade after the founding of NNIP, many of the parcel- and neighborhood-level data systems were funded in part by the Technology Opportunities Program (TOP) administered by the National Telecommunications and Information Administration of the US Department of Commerce. From 1994 to 2004, the TOP made matching grants to government agencies and community-based nonprofit organizations to demonstrate how digital networks could support community improvement. The National Telecommunications and Information Administration sought projects specifically related to local data intermediaries that fostered communication, resource sharing, and economic development within communities. Over these 10 years, 11 grants totaling $5.7 million were awarded to develop neighborhood-level information systems in seven cities.[18] TOP is a stellar example of how the federal government can accelerate innovation in support of the broader use of local data for decisionmaking.

One of the TOP-supported sites was Neighborhood Knowledge California (NKCA), developed by the Advanced Policy Institute at the University of California, Los Angeles, which also built its city's early warning system. Launched in 2003, NKCA included a variety of neighborhood-level national and local data for the entire state of California, including Decennial Census and HMDA data (Steins 2003). In addition to providing standard profiles, the site enabled users to generate their own charts and maps, to build their own neighborhoods from multiple census tracts, and to upload their local data for automated geocoding and display.

The Fannie Mae Foundation, which supported three of the four local early warning systems mentioned above, recognized that the NKCA efficiencies of assembling federal data for one state could just as easily be extended to the nation. Beginning in 2004, the foundation began the in-house development of the first national community indicators portal, and the website DataPlace went live the following year. Populated with several public datasets with small-area data prepared by the Urban Institute, the free site allowed users to explore the data through dynamic community profiles, tables, charts, and maps.[19]

In 2007, TRF created a similar national website called PolicyMap.[20] Still actively updated as of 2014, PolicyMap offers free access to indicators at the point, tract, and Zip Code levels from many federal datasets, as well as additional features and commercial data for subscribers. Other systems have developed to serve different audiences. For example, CommunityCommons.org includes most of the national datasets in PolicyMap and similar mapping functionality, but it emphasizes using data for health assessments and improvement initiatives.

A review of neighborhood information systems used qualitative and quantitative data from 2006 to 2007 to investigate the success of local information websites (Hwang and Hoffman 2009). Based on a survey of users of neighborhood information systems from around the country, the researchers highlighted the types of uses of the systems and the perceptions of usefulness. The current website picture for local sites is mixed.[21] Most of the local NNIP organizations have taken advantage of sophisticated data visualization and distribution systems to further their mission of democratizing data. By 2013 about half the NNIP partners hosted websites with interactive platforms. Some partners built in-house custom software, such as Cleveland's Northeast Ohio Community and Neighborhood Data for Organizing (NEO CANDO) system, which is featured in Nelson's essay at the end of chapter 5. Other partners chose to adopt commercial products from firms such as Azavea and Instant Atlas.[22] Still others have a simple descriptive website without data displays or tools. Although limited funding may contribute to the last choice, some organizations choose to focus their efforts on hands-on technical assistance and sharing data with targeted audiences, such as residents in low-income neighborhoods, who are not likely to use complex online systems.

A growing community of developer proponents and philanthropic efforts has fostered the open source software movement. In 2008, several NNIP partners and other data-related organizations formed the Open Indicators Consortium to develop open source visualization and data analysis software, later named Weave. The University of Massachusetts, Lowell, the consortium development lead, posted the Weave code in 2011, and the consortium grew to 16 local organizations by 2013.[23]

The Knight Foundation Community Challenges has spurred dozens of other applications to improve communities' access to data and news and expand the access to tools to connect with each other and tell stories using the data. Any product supported by these grants is required to be open source, increasing the chances of sustaining the product beyond the grant period and of reuse by other communities.

Other nongovernmental websites that host and visualize data have proliferated since 2010. The University of Minnesota houses the National Historical Geographic Information System to serve primarily academic researchers.[24] Funded by the National Science Foundation and National Institutes of Health, this system offers a free query system for Decennial Census, ACS, and Zip Business Patterns, with plans to add new datasets in the future.

### *Government Portals*

The advance of the government data sites paralleled the trend in nongovernmental sites. As early as 1997, the government launched FedStats in an attempt to create a one-stop shop for federal government data (US Census Bureau 1997). By 2002, the federal government distributed Decennial Censuses from 1990 and 2000, the Economic Census, and the pilot ACS through American FactFinder. This query system represented a leap forward from accessing data from file transfer protocol (FTP) sites or dozens of CDs, a step particularly critical to expanding access to users with less technical experience. Since 2005, OnTheMap has featured an extraction tool for its employment data, as well as mapping and reporting capabilities. Both American FactFinder and OnTheMap have been updated several times since their initial launch.

Open government data portals represent a major development in the evolution of data delivery by governments. The primary purpose of these portals is to publish entire data files, but some also have tools for visualization and sharing. Two main providers of open data portals are Socrata and CKAN. Most of the major city portals as of 2013 use Socrata, proprietary software that incorporates visualization and query tools. The Open Knowledge Foundation operates the open source CKAN platform, which is gaining in popularity. For example, data.gov was migrated from its original launch on Socrata to a CKAN-based system called the Open Government Platform in May 2013. The Open Government Platform is a joint open source product from India and the United States to promote transparency and greater citizen engagement by making more government data, documents, tools, and processes publicly available.[25]

With hundreds or thousands of data files, open data sites can be overwhelming, but other government agencies have developed websites with curated data and step-by-step guidance. These decision-support tools are aimed at specific audiences. They combine multiple sources of data relevant to a given task and provide visuals and guidance on using the data. An example is HUD's Office of Community Planning and Development's new eCon Planning Suite, which is defined as a "collection of new online tools to help grantees create market-driven, leveraged housing and community development plans" (US Department of Housing and Urban Development 2013b, 4). All localities that want to receive HUD grants under a number of programs are required to submit a consolidated plan. The new online template saves grantees time by prepopulating the plan with the latest housing and economic data related to their priorities. In addition, the eCon Planning Suite Maps tool has datasets available to help grantees assess market conditions for any user-defined geography. Grantees can insert maps and data tables they generate into their plan document.

### *Commercial Portals*

Many commercial firms serve the private sector by providing tailored tools to access small-area data for site selection and marketing. Some firms have also sponsored public websites as both a public service and as a marketing vehicle for their company and products. One early example is GeoCommons, which was unveiled in 2007 by GeoIQ, a private company that sells software and support to GIS enterprise customers. Users can use the GeoCommons site to explore, create, and share geographic data and maps licensed under a Creative Commons license. It was relaunched as a cloud-based service in 2011 with advanced analysis and development features (Gorman 2011). The site began with roughly 1,500 datasets at launch and grew to 50,000 open source geographic datasets in 2013 (GeoIQ 2007).

Google leveraged its expertise in search engines and user interfaces to create the Public Data Explorer in 2010. The site's goal is to help users find reliable data and statistics on a variety of subjects while providing tools to create dynamic mashups using graphs, maps, and bubble charts. The company prioritized what data it made available based on an analysis of the most popular data and statistics search topics. It began with data from the World Bank, the US Bureau of Labor Statistics, and the US Census Bureau, and later added five additional core data providers: the Organisation for Economic Co-Operation and Development, the California Department of Education, Eurostat, the US Centers for Disease Control, and the US Bureau of Economic Analysis.

Other notable examples of data and visualization sites are Tableau Public and Esri community maps, released in 2010, and IBM's City Forward, released in 2011. With the pace of technology advances, new easier-to-use and powerful websites will continue to emerge.

### *From Portals to Application Programming Interfaces and Mobile Applications*

With the expansion of mobile device use and capabilities, organizations from all sectors have expanded their focus from their own websites to Application Programming Interfaces (APIs) and mobile applications. Sanders's essay describes the mashups that are fueled by APIs, a protocol for requesting and delivering a data stream. In 2013, the Census Bureau launched APIs for the Decennial Census and the ACS, making it easier for developers to build applications using the data. Although mobile applications are now commonly used as a tool for data collection, there is even more potential for apps to display community data. Community-focused code-a-thons are time-limited events at which neighborhood and government representatives meet with software developers to identify community issues that could benefit from data visualized through an app. The field is advancing in how to structure these events to have concrete results, in addition to building bridges among the various groups.

## Coping with the Fast Pace of Progress

This chapter documents the remarkable progress that has been made in the availability of data and technology since the mid-1990s. Challenges, including those noted in Sanders's essay, remain for these advances to translate to better policy and practice. However, the field is learning what supports are needed to take advantage of the new opportunities. Chapter 2 discusses how NNIP partner organizations play this role in their local areas, but new national institutions are also emerging to strengthen the use of data in local public and social enterprises. DataKind, an organization founded in 2011, brings together leading data scientists with community organizations to provide help with data collection, display, and analysis for "positive action through data in the service of humanity." The organization's work ranges from short-term weekend "DataDives" to up to six months' placement of a technical staff person in the nonprofit. This type of community training and technical assistance has also been delivered by local actors. Chapter 5 discusses the evolution of the use of data in community development efforts in Chicago in Testing the Model and the collective impact perspective of the role of data in initiatives such as Strive.

Even with the new technological environment, state and local governments still grapple with legacy computer systems programmed in antiquated languages. In these cases, accessing the data remains time-consuming and cumbersome, sometimes controlled by a limited number of gatekeepers. The systems prevent either government agency staff or residents from accessing the data for improving operations or setting public priorities. Code for America was founded in 2009 to leverage private-sector energy and innovation to help address this issue. Code for America seeks to introduce advanced skills and technology into local government by pairing top technologists with local governments for one year to complete an open source civic technology project. Examples include a system in Detroit for collecting data from observational surveys and a website in New Orleans, Louisiana, to allow residents to view the city status of blighted properties. The time-limited effort has the potential to result in broader impact in three ways. First, Code for America leaders report that their one-year program often catalyzes new thinking in local government partners that extends to other technology decisions. Second, the products in many cases are being spun off to independent companies in hopes of spreading the innovation to other cities. Third, the Code for America efforts have tried to shift innovative technologists' image of government as a stagnant, hopeless bureaucracy to an important institution that provides vital services and is worthy of their professional time.

Finally, nonprofits and governments have a responsibility to recognize and mitigate the differential access to the new wealth of data and technology. In some ways the digital divide has narrowed. As documented by a Pew Research Center report, about 70 percent of African Americans used the Internet in 2011, double the share in 2000. Latinos showed a similar pattern of increase, with 68 percent accessing the Internet. These rates still lag behind the 80 percent rate for the white population, and the people left behind are less educated, poorer, and non-English speaking (Zickuhr and Smith 2012). Distressed neighborhoods are still going to be less connected to information than their wealthy counterparts. Living Cities recently commissioned a scan of civic technology to learn about the hurdles to having this new resource help low-income people. They found communities have challenges in creating tools that reflect community needs, navigating data privacy issues, allocating funding from tight budgets, and accessing people with high-tech skills (Hebbert and Ashlock 2012). The authors end with some suggested approaches to help civic technologists help low-income families and neighborhoods more often and at a deeper level. By harnessing the tremendous energy of the public and private organizations featured in this chapter, the field can improve its ability to connect information and technology to local decision-making and community improvement. Chapters 5 and 6 offer some inspiring local examples, and chapter 8 proposes a program of work to support these efforts nationally.

## NOTES

1. For an example from New Orleans, see "Residential Addresses Actively Receiving Mail by ZIP Code and Parish for the New Orleans Metro Area" (January 15, 2014), at the Greater New Orleans Community Data Center (http://www.gnocdc.org/ActiveResidentialAddresses/index.html).

2. For more information about the ACS, see *A Compass for Understanding and Using American Community Survey Data: What Researchers Need to Know* by Warren A. Brown (2009) at http://www.census.gov/acs/www/Downloads/handbooks/ACSResearch.pdf.

3. See Coulton (1999b) on vital statistics, Coulton (1999a) on welfare records, and Pettit et al. (2014b) for property records.

4. Other NNIP cross-site projects pioneered data development in health (Pettit, Kingsley, and Coulton 2003), foreclosures (Kingsley, Pettit, and Hendey 2009), and school readiness (Kingsley and Hendey 2010).

5. For other examples of parcel data, see *Transforming Community Development with Land Information Systems* by Sarah Treuhaft and G. Thomas Kingsley (2008) at http://www.lincolninst.edu/pubs/1356_Transforming-Community-Development-with-Land-Information-Systems.

6. For other examples of primary data collection, see the Memphis Neighborhood-by-Neighbor project by Tk Buchanan, Phyllis G. Betts, Jackson Gilman, Robert Brimhall (2010); the Project on Human Development in Chicago Neighborhoods at http://www.icpsr.umich.edu/icpsrweb/PHDCN/; and the tools at the Active Living Research website at http://activelivingresearch.org/toolsandresources/toolsandmeasures.

7. For additional information, see https://webfiles.uci.edu/kday/public/index.html.

8. For guidance on conducting community surveys, see Chris Hampton and Marcelo Vilela's "Conducting Surveys" at The Community Toolbox (http://ctb.ku.edu/en/tablecontents/sub_section_main_1048.aspx). Comey et al. (2013) also provides an overview of survey data collection and criteria for selecting a reputable survey research firm and issues to consider when designing and conducting a neighborhood or school climate survey.

9. To learn more about the research findings from the Project on Human Development in Chicago Neighborhoods, see Great American City: Chicago and the Enduring Neighborhood Effect by Robert J. Sampson (2012).

10. For an example of research using the L.A. Fans survey, see Clark (2012).

11. For additional information, see The Reinvestment Fund, "Limited Supermarket Access (LSA) Analysis Mapping Tool," at http://www.trfund.com/TRF-LSA-widget.html.

12. See Tauberer (2012) for an in-depth history of open data.

13. For additional information, see Philip Ashlock's blog post, "International Coordination for Local Collaboration" (March 5, 2010), at http://open311.org/blog/.

14. For additional information, see Google Developers at https://developers.google.com/transit/gtfs/.

15. For additional information, see http://housefacts.me/.

16. One of the hosting institutions, the Advanced Policy Institute at the University of California, Los Angeles, joined NNIP in 2002. The institute was reorganized as UCLA's Center for Neighborhood Knowledge and closed in 2009. Another host, the Center for Urban and Regional Affairs at the University of Minnesota, joined NNIP in 2007. A description of the Minnesota Neighborhood Information System is included in the chapter 2 essay by Wascalus and Matson.

17. The Chicago Neighborhood Early Warning System became Chicago's Community Information Technology and Neighborhood Early Warning System at. http://www.newschicago.org/about.php, but the data have not being updated since 2003 due to funding.

18. These cities include Chicago (two organizations), Indianapolis, Los Angeles, Minneapolis, New Orleans, and New York.

19. After the closing of the Fannie Mae Foundation, DataPlace was spun off to an independent nonprofit in 2007, which subsequently closed in 2009. The Urban Institute is no longer associated with the site.

20. For additional information, see http://www.policymap.org.

21. For a mid-decade review of the data delivery systems in Milwaukee, New Orleans, Providence, and Los Angeles, see Treuhaft et al. (2007).

22. See the NNIP Website Scan at http://www.neighborhoodindicators.org/library/catalog/nnip-web-site-scan-july-2014 (2014) for specific examples.

23. For more information, see the Open Indicators Consortium website at http://www.openindicators.org/portal. See also "Weave Beta 1.0 Announcement" at http://info.oicweave.org/projects/weave/wiki/Weave_Beta_10_Announcement.

24. The site was developed with the university's experience with their well-regarded Integrated Public Use Microdata Series site, which focuses on distribution of microdata.

For additional information about the National Historical Geographic Information System, see http://www.nhgis.org.

25. For additional information about the Open Government Platform, see http://opengovplatform.org/.

## REFERENCES


Beaghen, Michael, Tucker McElroy, Lynn Weidman, Mark Asiala, and Alfredo Navarro. 2012. *Interpretation and Use of American Community Survey Multiyear Estimates.* Research Report Series (Statistics #2012-03). Washington, DC: US Census Bureau.

Califano, Catherine, Kennen Gross, Lance Loethen, Scott Haag, and Ira Goldstein. 2011. *Searching for Markets: The Geography of Inequitable Access to Healthy and Affordable Food in the United States.* Washington, DC: US Department of the Treasury.

Carr, L. J., S. I. Dunsiger, and B. H. Marcus. 2011. "Validation of Walk Score for Estimating Access to Walkable Amenities." *British Journal of Sports Medicine* 45 (14): 1144–48.

Center for Transit-Oriented Development and Center for Neighborhood Technology. 2006. *The Affordability Index: A New Tool for Measuring the True Affordability of a Housing Choice.* Washington, DC: Brookings Institution.

Chan, Sewin, Michael Gedal, Vicki Been, and Andrew F. Haughwout. 2011. *The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City.* NYU Law and Economics Research Paper No. 11-36. New York, New York: New York University.

Chandler, Arnold, G. Thomas Kingsley, Josh Kirschenbaum, and Kathryn L.S. Pettit. 2006. The Potential of Parcel-Based GIS in Community Development and Urban Land Management. Boston, Massachusetts: Lincoln Institute of Land Policy.

Clark, William A.V. 2012. "Moving and Staying in Los Angeles Neighborhoods: Money Matters but So Does Family Composition." *CityScape: A Journal of Policy Development and Research* 14(3): 115–136

Comey, Jennifer, Peter A. Tatian, Lesley Freiman, Mary Kopczynski Winkler, Christopher Hayes, Kaitlin Franks, and Reed Jordan. 2013. *Measuring Performance: A Guidance Document for Promise Neighborhoods on Collecting Data and Reporting Results.* Washington, DC: Urban Institute.

Coulton, Claudia J. 1999a. *Public Assistance Records: A Source for Neighborhood Indicators.* National Neighborhood Indicators Partnership Report. Washington, DC: Urban Institute.

———. 1999b. *Vital Records: A Source for Neighborhood Indicators.* National Neighborhood Indicators Partnership Report. Washington, DC: Urban Institute.

Coulton, Claudia. 2008. *Catalog of Administrative Data Sources for Neighborhood Indicators.* Washington, DC: Urban Institute.

Day, Kristen, Marlon Boarnet, Mariela Alfonzo, and Ann Forsyth. 2006. "The Irvine–Minnesota Inventory to Measure Built Environments: Development." *American Journal of Preventive Medicine* 30 (2): 144–52.

Detroit Blight Removal Task Force. 2014. Detroit Blight Removal Task Force Plan. Accessed July 6, 2014. http://report.timetoendblight.org/index.html

Dowell, Earlene K. P. n.d. Local Employment Dynamics: Connecting Workers and Jobs. Accessed July 6, 2014. http://lehd.ces.census.gov/doc/learning_guide/LED_Basic_Training_2012-09.pdf

Duncan, Dustin T., Jared Aldstadt, John Whalen, Steven J. Melly, and Steven L. Gortmaker. 2011. "Validation of Walk Score for Estimating Neighborhood Walkability: An Analysis of Four US Metropolitan Areas." *International Journal of Environmental Research and Public Health* 8:4160–79.

Earls, Felton J., Stephen W. Raudenbush, Jeanne Brooks-Gunn, and Robert J. Sampson. 1995. *Project on Human Development in Chicago Neighborhoods: Community Survey, 1994–1995.* Ann Arbor, Michigan: Inter-university Consortium for Political and Social Research.

Environmental Protection Agency. n.d. Measuring Urban Heat Islands. Accessed July 6, 2014. http://www.epa.gov/heatisland/about/measuring.htm

Fellowes, Matt, and Mia Mabanta. 2008. *Banking on Wealth: America's New Retail Banking Infrastructure and Its Wealth-Building Potential.* Washington, DC: Brookings Institution.

Galster, George, Kathryn Pettit, Peter A. Tatian, Anna M. Santiago, and Sandra J. Newman. 1999. *The Impacts of Supportive Housing on Neighborhoods and Neighbors in Denver.* Washington, DC: US Department of Housing and Urban Development.

Galster, George, and Peter A. Tatian. 2009. "Modeling Housing Appreciation Dynamics in Disadvantaged Neighborhoods." *Journal of Planning Education and Research* 29 (1): 7–22.

GeoIQ. 2007. *GeoCommons Public Repository Now Includes over 2,500 Data Sets.* Accessed July 5, 2014. http://www.directionsmag.com/pressreleases/geocommons-public-repository-now-includes-over-2500-data-sets/112877

Goerge, Robert M. 2011. *Chapin Hall's Integrated Database on Child and Family Programs: Integrating Health and Human Services and Education Data to Inform Public Policy.* Presentation to Actionable Intelligence for Social Policy Fall 2011 Conference. Washington, DC. November 28–29.

Gorman, Sean. 2011. GeoCommons 2.0. Accessed July 5, 2014. http://blog.dc.esri.com/2011/05/31/geocommons-2-0/

Grover, Michael, Laura Smith, and Richard M. Todd. 2008. "Targeting Foreclosure Interventions: An Analysis of Neighborhood Characteristics Associated with High Foreclosure Rates in Two Minnesota Counties". *Journal of Economics and Business* 60 (1-2): 91–108.

Hebbert, Frank, and Phil Ashlock. 2012. *Field Scan of Civic Technology.* New York: Living Cities. November.

Hefter, Steven P. 2010. American Community Survey Sample Size Research. Section on Survey Research Methods—Joint Statistical Methods. Washington, DC: U.S. Census Bureau. 1183-92.

Hendey, Leah, Claudia Coulton, and Tom Kingsley. 2013. *Connecting People and Place: Improving Communities through Integrated Data Systems.* Washington, DC: Urban Institute.

Hickey, Robert, Jeffrey Lubell, Peter Haas, and Stephanie Morse. 2012. "Losing Ground: The Struggle of Moderate-Income Households to Afford the Rising Costs of Housing and Transportation." http://www.nhc.org/media/files/LosingGround_10_2012.pdf.

Hwang, Sungsoo, and Mark Hoffman. 2009. *In Pursuit of the Effective Neighborhood Information System: User-Friendliness and Training. Government Information Quarterly,* 26 (1): 166–173.

Immergluck, Daniel, and Marti Wiles. 1999. *Two Steps Back: The Dual Mortgage Market, Predatory Lending, and the Undoing of Community Development.* Chicago, Illinois: The Woodstock Institute.

Kingsley, G. Thomas (ed.). 1999. *Building and Operating Neighborhood Indicator Systems: A Guidebook.* Washington, DC: The Urban Institute.

Kingsley, G. Thomas, and Leah Hendey. 2010. *Using Data to Promote Collaboration in Local School Readiness Systems.* Washington, DC: Urban Institute.

Kingsley, G. Thomas, and Kathryn L. S. Pettit. 2008. *Data and Decisions: Parcel-Level Information Changing the Way Business Gets Done.* Washington, DC: Brookings Institution.

Kingsley, G. Thomas, Kathryn L. S. Pettit, and Leah Hendey. 2009. *Addressing the Foreclosure Crisis: Action-Oriented Research in Three Cities.* Washington, DC: Urban Institute.

Kitzmiller, Erika. 2013. IDS Case Study Chapin Hall: Leveraging Chapin Hall's Mission to Enhance Child Well-Being. Philadelphia, Pennsylvania: University of Pennsylvania.

Klosterman, Richard E., and Alan A. Lew. 1992. "Tiger Products for Planning". *Journal of the American Planning Association* 58(3):379–85.

Lane, Julia, Pat Doyle, Laura Zayatz, and Jules Theeuwes, eds. 2001. *Confidentiality, Disclosure and Data Access: Theory and Practical Applications for Statistical Agencies.* North Holland.

LaVigne, Nancy G., Jake Cowan, and Diana Brazzell. 2006. *Mapping Prisoner Reentry: An Action Research Guidebook.* Washington, DC: Urban Institute.

MapBox. 2013. *OpenStreetMap Data Report.* Accessed July 7, 2014. https://www.mapbox.com/osm-data-report/

Mayer, Neil S., Peter A. Tatian, Kenneth Temkin, and Charles A. Calhoun. 2011a. *Has Foreclosure Counseling Helped Troubled Homeowners? Evidence from the Evaluation of the National Foreclosure Mitigation Counseling Program.* Washington, DC: Urban Institute.

Mayer, Neil S., Peter A. Tatian, Kenneth Temkin, and Charles A. Calhoun. 2011b. *National Foreclosure Mitigation Counseling Program Evaluation: Final Report Rounds 1 and 2.* Washington, DC: Urban Institute.

McCoy, Patricia A. 2007. "The Home Mortgage Disclosure Act: A Synopsis and Recent Legislative History." *Journal of Real Estate Research* 29 (4): 381–97.

National Neighborhood Indicators Partnership. 2014. "NNIP Data Inventory by Data Source." Washington, DC: Urban Institute.

Pettit, Kathryn L. S., and Audrey E. Droesch. 2008. A Guide to Home Mortgage Disclosure Act Data. Washington, DC: Fannie Mae Foundation.

Pettit, Kathryn L. S., Leah Hendey, Brianna Losoya, and G. Thomas Kingsley. 2014a. Putting *Open Data to Work for Communities.* Washington, DC: Urban Institute.

Pettit, Kathryn L. S., G. Thomas Kingsley, and Claudia J. Coulton. 2003. *Neighborhoods and Health: Building Evidence for Local Policy.* Washington, DC: US Department of Health and Human Services.

Pettit, Kathryn L. S., Barika X. Williams, G. Thomas Kingsley, Peter A. Tatian, and Leah Hendey. 2014b. *Using Property Databases for Community Action.* Washington, DC: Urban Institute.

Providence Plan. 2013. About the Data Hub. Accessed July 5, 2014. http://ridatahub.org/about/

RAND Corporation. n.d. LA FANS Documentation. Accessed July 5, 2014. http://lasurvey.rand.org/documentation/

Snow, Christopher, Kathryn L. S. Pettit, and Margery Austin Turner. 2003. Neighborhood Early Warning Systems: Four Cities' Experience and Implications for the District of Columbia.

Spielman, Seth E., David Folch, and Nicholas Nagle. 2014. "Patterns and Causes of Uncertainty in the American Community Survey." *Applied Geography* 46:147–157.

Steins, Chris. 2003. "E-Government and Planning Profiles: Neighborhood Knowledge California." American Planning Association. January. Accessed July 20, 2013. http://www.planetizen.com/node/9156.

Tauberer, Joshua. 2012. Open Government Data. Accessed July 5, 2014. http://opengovdata.io/

Treskon, Mark, Milena Kornil, and Josh Silver. 2003. *Broken Credit System: Discrimination and Unequal Access to Affordable Loans by Race and Age: Subprime Lending in Ten Large Metropolitan Areas.* Washington, DC: National Community Reinvestment Coalition.

Treuhaft, Sarah, Arnold Chandler, Josh Kirschenbaum, Melissa Magallanes, and Randal Pinkett. 2007. *Bridging the Innovation Divide: An Agenda for Disseminating Technology innovations within the Nonprofit Sector.* PolicyLink: Oakland, CA.

Urban Institute. 2013. Catalog of Established and Emergent Integrated Data Systems. Accessed July 6, 2014. http://www.neighborhoodindicators.org/library/catalog/catalog-established-and-emergent-integrated-data-systems.

Urban Land Institute. 2006. *The Beltway Burden: The Combined Cost of Housing and Transportation in the Greater Washington, DC Metropolitan Area.* Washington, DC: Terwilliger Center for Workforce Housing and the Center for Housing Policy, Urban Land Institute.

US Census Bureau. 1997. Press Release: Administration Introduces One-Stop Federal Statistics Web Site. May 22. http://www.fedstats.gov/osspress.html

US Census Bureau. n.d. About the ACS Compass Products. Accessed July 6, 2014. http://www.census.gov/acs/www/guidance_for_data_users/compass_products/

US Department of Housing and Urban Development. 2010a. *HUD's Fiscal Year (FY) 2010 NOFA for the Choice Neighborhoods Initiative—Round 1 NOFA.* http://portal.hud.gov/hudportal/documents/huddoc?id=DOC_9823.pdf

US Department of Housing and Urban Development. 2010b. *HUD's Methodology for Allocating the Funds for Neighborhood Stabilization Program 3 (NSP3).* http://www.huduser.org/portal/datasets/NSP3%20Methodology.pdf.

US Department of Housing and Urban Development. 2013a. Data and Methodology: Location Affordability Index and My Transportation Cost Calculator. Washington, DC: US Department of Housing and Urban Development.

US Department of Housing and Urban Development. 2013b. *The eCon Planning Suite: A Desk Guide for Using IDIS to Prepare the Consolidated Plan, Annual Action Plan, and CAPER/PER.* Washington, DC: US Department of Housing and Urban Development.

US Department of Housing and Urban Development. n.d.a. Homeless Management Information Systems (HMIS) website. Accessed August 13, 2013. http://portal.hud.gov/hudportal/HUD?src=/program_offices/comm_planning/homeless/hmis

US Department of Housing and Urban Development. n.d.b. NSP2 Data and Methodology. Accessed July 5, 2013. http://www.huduser.org/portal/NSP2datadesc.html.US Government Accountability Office. 2013. *Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy*. Washington, DC.

US Government Accountability Office. 2013. *Sustained and Coordinated Efforts Could Facilitate Data Sharing While Protecting Privacy.* Washington, DC.

Walk Score. n.d. "Case Study: Analyzing Light Rail Station Area Performance in Phoenix." Accessed October 20, 2013. http://cdn2.walk.sc/pdf/WS-Phoenix-TOD.pdf.

Zickuhr, Kathryn, and Aaron Smith. 2012. *Digital Differences.* Washington, DC: Pew Research Center.

ESSAY

# Mashup City

## *Tools for Urban Life and Urban Progress in the Internet Age*

Greg Sanders

In 2011 a Chicago, Illinois, technology firm launched a digital service that allowed mobile phone users to find nearby restaurants serving their favorite meals. Type *pasta primavera* into your phone, and you received not only a list of nearby offerings, but ratings—and not only for each restaurant, but for each restaurant's pasta primavera. It was just another reliable, customized, granular data product served up for mobile consumers in a matter of milliseconds. Nothing remarkable there. Since restaurant locators without menu-level detail were already popular, of course someone would improve and extend those services within a year or two. Our expectations for more detailed, localized, personalized information, growing on an ever-steepening upward curve, were confirmed and raised a little bit more.

Meanwhile, across town, a community revitalization initiative brought together government, nonprofit, university, and volunteer groups for a collaborative effort to improve the neighborhood. They planned a commercial corridor renovation with mixed-income residential units and green space. But they needed data, and lots of them. They needed to find out who owned all 800 parcels within the project area. They sought data on zoning, land use, and property tax status; traffic counts; crime statistics; and income and spending. They needed historical numbers to effectively tell the story of how their community had changed, for better and worse. And they needed to sort, query, visualize, and analyze all that data. But while enthusiasm for the project grew, data acquisition and analysis efforts bogged down and came to a halt. Data couldn't be found at the necessary detail or at all. Some resources remained in a paper format; others were available only at a prohibitive cost. Most of the data needed to be entered into a desktop database before it would be usable. Expectations for success sank lower and lower.

Does it need to be this way? Will software providers always prefer marketing products for foodies and social media addicts, while mission-driven devotees of social progress languish in a technology desert?

After describing the need for and traditional barriers to information in community efforts, this essay shares some reasons for optimism. The first bit of good news is that the raw data needed to create community revitalization tools are more available today than at any time in our nation's history. The second bit of good news is that several technologies behind the recent explosion of data-driven apps are nonproprietary and can be leveraged by anyone with the know-how to use them. The combination of abundant raw materials and low-cost tools could greatly benefit community revitalization practitioners. So what will it take to bring forth a new set of data-driven community revitalization tools? The final sections suggest some answers to that question.

## The Need for Information in Community Revitalization Efforts

If community revitalization is about improving quality of life, information technology is about improving the quality of action. Action based on good information should be more effective, more efficient, more targeted with fewer unintended consequences. Well-informed actions should yield better quality-of-life improvements than poorly informed ones.

Data systems add value to community revitalization efforts for two distinct groups: individual users and institutional users. For community revitalization to succeed, individuals need access to information that extends opportunities to them. But community success also depends on institutional intermediaries, whose opportunities focus on land use, public safety, economic development, and transportation. This premise of organized community revitalization work is that information does not necessarily need to be consumed by the ultimate beneficiaries (the residents, employees, or visitors in the community). Information consumed by advocates can yield results that extend to the whole community, or at least some of its members.

This distinction is good to keep in mind when considering the "by whom, for whom" aspects of information. Community members can make good use of some kinds of information, while advocates can make better use of other kinds, even though the intended beneficiary is always the community (or so we would hope).

The crux of information's value lies in the discovery and assessment of opportunity. Community residents discover free workshops or medical screening through information systems and, if their assessment of the opportunity is positive, they avail themselves of it and reap the benefits. Advocates might discover vacant and available properties that can be converted to community assets if their assessment of the property's value and suitability are good. Barriers to opportunity are by no means all information related, but lack of discovery is surely the first barrier. After discovery comes an information-based assessment of the opportunity, followed by information-based activation of the results. It does not help to have a medical screening indicating a need for treatment if there's no information about how to get treatment. Advocates who locate a property suitable for development can do little if they can't identify the property owner.

Noninformational barriers such as a lack of material resources might far outweigh informational barriers, but there can be no doubt that without information, opportunity is always elusive. Better-informed actors always have an advantage over poorly informed actors (though not always a decisive advantage).

## Barriers to Community Access to Information

Despite the need, local community improvement efforts rarely have the data or tools they need to make informed decisions. The *digital divide* normally describes the lack of affordable access to internet bandwidth and other digital tools. But this term might also describe the chasm between technologies targeted at mass market consumers and technologies intended for nonprofit community revitalization work. The market for community-focused data products is not insignificant, but it is driven mainly by resource-poor customers. The result is an inadequate supply of affordable tech tools for the important work of community and economic development.

Several barriers have historically kept community activists from the data promised land:

- *Government as fortress.* Governments at all levels have traditionally acted as protectors of information, charged with keeping their secrets out of hostile hands. Careers have been wrecked over data breaches when hackers have exposed confidential information. Unsurprisingly, government workers have responded by circling the wagons and making data protection their highest priority, even for nonconfidential information.
- *Intellectual property rights.* Many important datasets are readily available to community activists, sort of. The "sort of" qualifier can take many forms. Some datasets are available for a fee; others are available at no charge but cannot be freely distributed. Some require laborious negotiations over terms of use, which might drag on for months or years.
- *Complexity.* Information systems are almost always built on complex architectures. Technicians routinely speak of "hiding" complexity. Grady Booch, a guru of object-oriented programming, famously declared that "the function of good software is to make the complex appear to be simple" (Fitzek et al. 2010)—which is great unless you're the person in charge of simplifying the complex, in which case your job is anything but simple. Community-based intermediaries often must accomplish the simplification, which has proved to be a formidable stumbling block.
- *Diverse needs.* Activists find themselves in the position of needing to fill a daunting variety of information needs, ranging from the personal needs of individual residents to the institutional needs of community organizations. Such a wide divergence of required solutions can be crippling, because the solutions might have little overlapping functionality. Code reuse across systems is always ideal, but there are inherent challenges in designing information systems that serve such a range of needs.

The remainder of this essay describes the progress to date in overcoming these barriers, as well as prospects for the future.

## Wholesale and Retail Information Markets

For information to be accessible for community improvement, both the wholesale and retail information markets must be healthy. Wholesale datasets—large-scale numeric or unstructured text files—are seldom useful without further processing. Retail information—maps, charts, animations, easily consumable tables, and other user-friendly outputs—are much more valuable to the end consumer. Conversion of any raw material into finished products is typically handled by value-adding artisans and manufacturers. Data resources work in the same way: when raw data are available and demand for finished products is strong, developers will emerge to create the finished products. When raw data are unavailable or the demand for value-added products falls under the threshold of profitability, the raw-to-ready conversion never gains critical mass.

Until recently, data tools that were relevant to community revitalization efforts were doubly doomed. Not only were financial resources (i.e., demand for finished data products) lacking, but the raw data needed to manufacture them were scarce. Recently the picture has brightened for both markets.

### Progress in Wholesale Information Markets

Government as fortress was the norm in the United States for two centuries before the digital age began to tear down small chunks of the fortress walls. Soon after the Internet became a prominent feature of everyday life in the 1990s, public agencies at all levels began posting information in static Hypertext Markup Language (HTML) and Portable Document Format (PDF) documents. These documents needed to be downloaded one at a time and converted into dynamic database formats that could be queried. Screen scraping (computer programming that automated the download of HTML text) was commonplace but suffered fatal breakdowns every time the HTML documents were structurally revised. Within a few years, downloadable spreadsheets and dBase files became available from the US Environmental Protection Agency, many state police agencies, and other agencies. But the first real breach in Fort Government occurred with the advent of protocols for retrieving machine-readable data. (*Machine readable* means a computer can comprehend and manipulate data without human intervention and can produce whatever the user has requested, such as maps, charts, or reports.)

Machine-readable data could be requested in filtered and sorted formats directly from the source in real time and displayed for the end user in many formats in any web browser.

In 2003 Phil Windley, the chief information officer for the State of Utah, published his "Web Services Manifesto," arguing that all data resources should produce at least machine-readable Extensible Markup Language (XML) data, and human-readable text only when necessary. The idea was to let others develop applications built on the machine-readable government data. Windley wasn't the only voice calling for government data to be made available as open web services, but his formulation was elegant and actionable. In 2006, we in Chicago formed a data exchange collaborative (the Illinois Data Exchange Affiliates) around Windley's principles (Sanders 2007). In 2007 a group of open government advocates developed the "8 Principles of Open Government Data" (Open Government Working Group 2007).

On his first day as president in 2009, Barack Obama issued a Memorandum on Transparency and Open Government (The White House 2009) that not only set the tone for federal policy on open data, but also sent a message to state and local government entities that might seek federal funding. By 2009 the federal government had launched the http://data.gov open data portal. Progress is not limited to the United States. Many European countries (and the European Union itself) have rolled out open-data platforms. After the launch of data.gov, publisher and web guru Tim O'Reilly popularized the term *government as platform,* which conceptualized the public sector as a provider of raw information from which the private sector could build applications.

The federal government continues to increase its annual data output, and many of its wholesale products are directly relevant to community revitalization work. Data from the US Census Bureau have long been packaged by entrepreneurs into many useful tools, but the Environmental Protection Agency and the Departments of Labor, Education, Transportation, Housing and Community Development and others also contribute to this flow.

Data access below the federal level has improved also. Many state agencies have released granular data about issues of community importance such as public health, education, and licensing. Local government, long associated with impenetrable data silos, has made good progress in some cities, notably New York, Chicago, San Francisco, and Washington, DC. Early proof of the government as platform concept emerged in 2005 in the form of chicagocrime.org, a forerunner of EveryBlock.com (which was created a few years later by the same developer, Adrian Holovaty). But even in progressive regions, municipal governments outside the core city almost universally remain impenetrable data fortresses. Despite the fact that many local governments have not jumped aboard the open data train, it brings hope to all of us who remember a time when almost no noncensus data were digitally available at the neighborhood level.

Of course, government does not bear the whole burden of making good data available. Corporate, academic, and civic institutions can play an enormous role in releasing high-value data for community improvement. Progress toward open data has been slower in these non-governmental sectors, as private intellectual property and corporate trade secrets aren't covered under the Freedom of Information Act. For example, telecommunications companies have long resisted releasing data about the number, type, and location of their broadband internet customers, just as energy distributors resist releasing energy consumption data. Universities and nonprofits are often understandably reluctant to release their value-added data products, usually the result of extensive labor. But in the age of open data, it might be more beneficial for these providers to release their works for free use and accept the credit, rather than try to monetize the resource.

## Progress in Retail Information Markets

Although American government's great leap forward into open data has placed abundant resources in the hands of community revitalization practitioners, abundant raw data do not immediately translate into usable information. To the extent that government is in the information business, it deals in wholesale, not retail assets. That is, it produces very few finished products that constituents can consume directly. And its products tend to be less useful at the more granular levels that are relevant to local communities (e.g., census blocks, parcels, points, and street lines).

But there's a ray of hope. In addition to the increased availability of raw data, new technologies can enable the creation of data processing and visualizations at a low cost for high-skill technicians. Leveraging hard skills to reduce the cost of critical products and services is nothing new. Think about the cost of wiring and plumbing an older home as a do-it-yourself project versus contracting the entire job. Expertise pays off not only in labor markets, but in private life as well. The home handyperson possesses a concrete and tangible resource that can make the difference between affordable transformation of available materials and limited or no transformation. Without expertise, one's options are to live with the old plumbing and wiring or incur unaffordable renovation contract costs.

Civic hackers have expertise in the new technologies and an interest in serving the public. They leverage raw data to produce high-quality maps, charts, and other outputs. Code for America has led the charge to enlist talented developers and designers to develop web applications for city governments. These new applications have been made possible by the continuous simplification of tools for converting raw data into end-user data tools. Open source platforms and code-sharing repositories have led to a rapid expansion of software for quickly generating maps and charts from raw data. The following section describes some of the key components of this new technology.

## New Enabling Technologies

The not-quite-fictitious food finder app described above was made possible by the same technologies that drive today's dizzying array of consumer applications. At the time of this writing, the salient technologies are mashups, application programming interfaces (APIs), mobile, geolocation, cloud, and search. Many other elements of traditional information systems remain important today, such as transfer protocols, disk storage, and database tools. But the recent consumer applications, while incorporating that older foundation, get their wings from these "big six" tech-architecture newcomers.

These technologies are most easily described in the context of popular consumer products; we'll return to community revitalization after laying out the basic tech terrain in place at this writing. The essence of today's tech stack is this: Pull together as many pieces as the consumer needs, from wherever those pieces are stored, and deliver them to the consumer based on highly customized search criteria. And do it fast. So let's talk about the star of this show, the mighty mashup.

A mashup is a combination of multiple data resources to form a more useful end product. For example, a mashup might pull together a list of restaurants located in the customer's vicinity, then combine it with a database of that customer's food preferences. One service provides the list of nearby restaurants, and the user's preferences are pulled from another. It matters not at all whether the provider of restaurant data uses completely different tech platforms than the provider of food preference data. All that matters is that they both quickly deliver data in a machine-readable format after submitting a valid service request. The most famous mashup factory in existence today is the ubiquitous Google Map, which combines the customer's personal data—a file in Keyhole Markup Language (KML) format on a local device—with Google's Map engine.

Mashups work because data providers publish APIs. An API is essentially a contract to deliver a specified data product in a specified format in response to a valid request. Some APIs are freely accessible to all, while others come with a cost. APIs work because the Internet works. That is, HTTP (hypertext transfer protocol) facilitates the transport of a relatively large amount of information over a wired or wireless network. Request in, data out. The network is the highway; HTTP is the transport company. APIs accept the customer's order and ship the goods. It's up to the receiver to do something with those goods. Twitter's API is a good example. If you want to find all tweets marked with the hashtag #*nnip*, you can call the Twitter API and receive a list of those tweets.

But you probably wouldn't do that unless you were an experienced web developer. More likely, you'd open your mobile phone and use an app created by a web developer. You would type #*nnip* and click *Go* and the list would appear. Raw data transformed by a high-skill technician appears as a finished product for the consumer.

The mashup's chain of events typically follows a pattern of this kind:

- The user navigates to a web page by using a browser.
- The web browser contacts a web server for the page content.
- The web server pulls data from a local database, but also requests additional data from a remotely hosted API.
- The user's browser displays the now-hybridized content (pulled from two or more servers), including clickable options for further exploring.
- With each click from the user, the browser itself initiates API requests for more data from remote servers.
- The new data are displayed in a section of the web page, but the web page itself does not need to reload when this occurs.
- The process can go on indefinitely without the user ever needing to reload the web page.

This kind of information tool would be great fun, and maybe valuable, even if it sat on your desktop computer. But you don't need a list of restaurants located in the vicinity of your desktop computer nearly as much as you need a list of restaurants near your hotel in a city far from home.

Mobile technology not only is more convenient than nonmobile, but it might also be the only available technology for some users. Low-income audiences might forego the landline phone and the desktop computer in favor of a mobile phone, whether it's a smart device or not. Mobile technologies rely not only on wireless/cellular networks, but increasingly on geolocation, a satellite-based system of determining a mobile device's (almost) exact position at any given time. Geolocation gives us the possibility of scoping all data requests to the device's vicinity. A provider of data about nearby restaurants knows what area to focus on without the inconvenience of customer-entered location data. This ability can be enormously important for collecting point-level data, such as the location of a 911 call or even a reported pothole. Users might have some notion of where they are at any given time, but they might express that information in any number of ways that are difficult to decipher. Geolocation can solve this problem.

When mashups request data from an API, they call a web address that probably exists in a cloud. Clouds come in many shapes and styles, but the essential nature of a cloud is that it combines physical assets (storage devices, processors, switches, and so on) in a converged architecture, and that almost always means virtualization. Virtualization creates "servers" on demand out of physical resources grabbed from farms of disks, processors, and memory. Cloud computing is important for today's tech systems because of cost efficiency and because clouds obviate the need for data providers to maintain on-site hardware.

All the technologies described here are made more powerful by search engines, which have improved immensely in recent years. Mashups are great, but they're less helpful if you don't know they exist, or if they don't have a search engine telling them what search terms you specified. Google is the current search leader, but search technology goes far beyond Google. The term can be applied to machine learning and other predictive methods that quickly lead users to the resources they seek.

Enabling technologies are just that—tools that enable some action or move us closer to a goal. But technologies don't live our lives for us, and they don't achieve our goals. We need to deploy the tools within the flow of our professional and personal lives. Tools can be deployed in the service of the mission of community revitalization, contingent on raw materials, tech expertise, and other resources. So the question remains: Why can lovers of pasta primavera instantly discover the best platter within walking distance, while community activists struggle to find and deploy basic information resources in the service of their cause?

## Making Information Actionable

The previous two sections document the improvements in access to data and technology that have enhanced the potential for community groups to access "retail" information tailored to their needs. Although this information can be a useful means to effective action, it is not in itself a useful outcome. Too often information is viewed as a result, rather than as a means to achieving a result. This misplaced focus on information for its own sake sometimes generates data systems that lack applicability to real-world problems. In fact, it's fair to say that all information is worthless to some people, and all information is valuable only within a given set of circumstances. In the context of community revitalization, the following factors determine whether information is useful and actionable:

- *Intelligibility.* Who can comprehend the information, and under what circumstances?
- *Accessibility.* Who can access the information, and who cannot?
- *Timeliness.* Is the information acquired at a time when it can be acted on?
- *Resources.* What is needed to convert information to effective action?

As these questions imply, the limits of information's effectiveness are strongly situational, varying from individual to individual and moment to moment. Given the situational, contingent nature of information and its usefulness, can there be any doubt that some people find themselves at a disadvantage in the game of leveraging information? A few examples can illustrate this point.

Failures in intelligibility start with language barriers, which can render even the clearest information worthless. Language conversion is already implemented in many data systems, but it is far from perfect. Some information systems convey information in a highly graphical way, but most do not. Educational achievement can significantly affect intelligibility. Information systems require a level of expertise ranging from very low to very high. The lower a person's expertise in the relevant subject, the fewer information resources can be leveraged. Illiteracy renders most information systems totally or nearly worthless.

Accessibility overlaps with intelligibility (especially in the realm of disabilities such as impaired vision), but it includes physical barriers to information media. The decades-long efforts to conquer the digital divide have had some important victories, but many people still lack access.

Timeliness is important because effective action is almost always bound by windows of opportunity. It doesn't help to learn about an opportunity after it has passed. In one tragic but instructive example, the massive public service campaign against cigarette smoking in the late 20th century came too late for many smokers, who were already ill or addicted when the information became widely available. In a more mundane example, a train- or bus-tracking system that broadcasts a service delay only after it has been resolved is not very useful. When individuals are required to invest time in overcoming barriers to information, they often acquire the information after it has lost some or all of its value.

Finally, information alone is rarely actionable, because most action requires both information and noninformation resources, such as money, time, transportation, and so on. Information about a subsidized loan program for homebuyers may require that the applicants be bankable, but it is not truly actionable in the absence of financial resources. Knowing that a clinic is offering free screening is less valuable to someone without transportation resources. Information about free public training workshops may have little value for a single parent working two jobs and thus lacking time.

## Opportunities for Incremental Change

Clearly, there is much room for improvement in the realm of information tools for community revitalization. The essential challenge is to lower barriers to software creation and raise our collective capacity for overcoming the barriers that remain. The following strategies for improvements in data, technology, and skills might seem obvious, but they deserve to be evaluated as potential opportunities for intervention.

## Invest in Data

### *Increase the Availability of Raw Data*

Despite the rapidly growing volume of open data published by government and nongovernmental sources, many more datasets remain siloed. But siloed information is not the only potential lode to be mined if we wish to increase the data supply. Great volumes of data could be generated by the simple process of aggregating raw data geographically (e.g., calculating sum and mean for the value of foreclosed properties at the city, county, and school district levels), chronologically (e.g., calculating mean values for each month, quarter, and year), or topically (e.g., combining small categories into broader ones). These aggregated datasets would still constitute a kind of raw data in that they would be numbers rather than maps, charts, or other retail products.

Nongovernmental data, too, might make a valuable contribution to the raw data supply. For example, researchers and advocates would love to have access to customer usage data from the broadband providers and energy distributors noted above, but currently they can only make educated guesses.

Increasing the already-massive volume of open data might seem like overkill, unless you are a stakeholder searching for granular data relevant to an issue that is important to you. This dynamic—that is, atomized needs for specialized data—continuously puts pressure on governments and other providers to release everything, all the time, in formats disaggregated geographically, chronologically, and topically. The cost of continuous data updates can be quite high, and the provider runs a significant risk of consumer backlash when expectations are not met for issues such as accuracy, formatting, frequency of update, and too much or too little protection of personally identifiable information.

### *Increase Data Quality*

Data quality encompasses everything from address cleaning to granularity to how the data are structured. But perhaps the most important aspect of data quality is the inclusion of elements that link each row of data to other resources. The best example of this is geocoding: Much government data are logically tied to a place (e.g., bridge, property, fire hydrant), and that location is all important for analysis. Does Wisconsin have more deteriorating bridges than Illinois? A geocode stamped on each bridge record would make it easy to answer that question. Nongeographic linking elements might include the license or serial number for physical assets like vehicles and computers. (Social security numbers are the ultimate linking elements, a fact that emphatically demonstrates the limits of open data from government sources—linking disparate data sources is good, except when it's bad. When privacy is at stake, linking informational elements is very, very bad.)

## Invest in Technology

### *Simplify Data Processing Tools*

Huge progress has already been made in simplifying data-processing tools, but still the conversion of raw data to finished data product is beyond the capacity of almost everyone. Google Maps notwithstanding, most data visualization software is still too complex to be useful. The good news is that many large and small software firms have a strong incentive to continue simplifying their tools. Capturing market share is not the only motivator here—open source communities and hobbyists are as eager as the bigger players to innovate in this process.

### *Make Simple Data Collection Tools Available*

Disseminating simple tools for informal data collection would help intermediaries to build data resources that are otherwise not available. Examples would be surveys of local residents, curbside surveys of physical conditions in the community, inventories of local employers, and green-space mapping. Although some tools are already available for this kind of data collection, they are not well-known or understood. Because such tools are most effective when customized for local conditions, the need for them is less likely to be satisfied by generic one-size-fits-all software solutions.

### *Share Code Broadly and Simply*

Code sharing has in the past decade become widespread, especially since the emergence of GitHub as a near-universal code repository. This trend should not be underestimated as a force in nurturing innovation among software developers. Code sharing is the essence of open source software development, which has proved to be an excellent catalyst for the creation of affordable software tools.

Encouraging code sharing and collaboration through local incubators (as Boston and other cities have recently done) is an effective path toward developing a strong community of local software technicians.

### *Improve Identity Management*

Social service providers, like health care and financial service providers, are burdened by a critical need for stringent protection of private information. But social service providers typically do not command the technological resources available to hospitals and banks, and therefore may be less able to create and maintain secure web systems. Many social service agencies gather detailed data about their constituents' problems—legal, financial, medical, and other issues. The potential for creating integrated systems to improve human service delivery has been articulated by thoughtful practitioners [see, for example, Allegheny County's "Building an Interoperable Human Services System" (Smith 2008)]. But truly integrated data systems with multiple distinct partners are difficult to lock down. As impressive as Allegheny County's achievement has been, that success would have been far more complex if it had linked multiple private providers rather than divisions of local government. The more independent nodes a network contains, the more difficult data protection becomes.

This integration problem is really an authentication problem. Authentication is the verification that data requesters are truly who they say they are, and that they have permissions to access the requested data. A secure system must fulfill 100 percent of valid requests by authorized users, and must deny 100 percent of requests by unauthorized users. So identity management services, which accept each user's identity claims and respond with a decision on the access request, are the linchpin of secure systems. Some promising standards for identity management have emerged in recent years.

Public key infrastructure offers a powerful identity management model, but it is currently more feasible in the context of corporate and government employee authorization management than widespread public use. oAuth is an ingenious solution used by web developers when providing personalized services to their end users, but it is not yet widely accepted for protecting truly confidential personal data. Investing in the further development of these models might improve our ability to serve diverse community audiences.

## Invest in People

### *Train (or Recruit) More Software Programmers*

For all the progress we hope to make toward software simplification, it's important to recognize that even the best software does not eliminate the need for competent data technicians. Increasing data processing and software engineering skills across the board is a challenging task, but the good news is that every successive cohort of school children will likely start middle school at a measurably higher skill level than its predecessors. This generational improvement could be accelerated through public investments in science, technology, engineering, and math education. As we broaden and deepen the community of technicians capable of turning raw data into retail data tools, opportunities for progress toward community revitalization will increase. Today, the demand for skilled software and data technicians still far outstrips the supply. Encouraging immigration of skilled technicians through the H1-B visa program can yield excellent results in this regard, but the federal government continues to resist expansion of this valuable program.

### *Promote Data Literacy*

Perhaps an even larger challenge is to raise the capacity of end consumers to leverage data tools for improvements in their quality of life. Improving people's ability to effectively leverage data tools is more challenging than improving the skill of our software technicians. Technicians have voluntarily sought to learn the skills critical for their craft, but the people who most need to leverage data tools are often unwilling or unable to learn how to do so. The pattern is familiar: Tools and resources often are rejected or ignored by those most in need of assistance. Such resistance may persist regardless of efforts to help, but some interventions, such as ubiquitous affordable broadband networks, increased investment in technology education, and face-to-face outreach, can at least set the preconditions for gradual improvement.

### *Promote Evidence-Based Decisionmaking*

Analogous to the problem of training individual end users to effectively use data tools is the problem of promoting data-driven decisionmaking among institutions. No amount of great data or software can have an impact on our communities if decisionmaking is not fundamentally tied to data analysis. Governments and community and economic development organizations have begun to weave data analysis into their processes, but this transformation is incomplete at best. Only when stakeholders (e.g., taxpayers) demand to see the evidence underlying decisions affecting the community will decisionmakers embrace data analysis as a standard operating procedure. All the opportunities discussed above include the potential for public policy to improve the effectiveness of technology in community revitalization. Governments should release more and better data; invest in technology education, outreach, and broadband availability; encourage innovation and collaboration to strengthen local tech communities; broaden the H1-B visa program; and promote a culture of evidence-based decisionmaking in the public sector. But community revitalization practitioners—and their constituents—must embrace and invest in technology expertise to reap the benefits of these increases in data availability.

## Opportunities for Fundamental Change

As the pace of change steadily increases from disruptive to frenetic, predicting the future becomes nearly impossible. We can only predict that change will continue, and guess at some directions in which innovation might blossom. But it is possible to speculate on the broad outlines of how innovation will progress, if not the direction it will take. One safe bet is that innovation will continue to progress along a path of interconnection among diverse, but related, elements. Social networking is an example of such interconnection, as are mashups. Google has connected its Gmail, Calendar, Docs and Google+ platforms into an integrated personal communication and productivity suite. But imagine if all information relevant to all aspects of and places within a particular community were linked automatically whenever any new information was added to any relevant system. This is the vision of the Semantic Web, defined by the World Wide Web consortium in this way: "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries." It requires that data systems include tags (including geographic identifiers) that point to known entities such as businesses, land parcels, and even human beings. Application developers who could tap into a full-fledged Semantic Web would have innumerable options for mashing up the information in valuable ways. Community and economic development could rely on extensive, frequently updated databases at all geographic levels, from parcel to county.

If identity management were strengthened and woven into the Semantic Web, individuals could link all aspects of their health, educational, legal, social, financial, logistical, and other data into—of course—value-added retail data tools. The holy grail in this case would not be an assembly of a person's known information resources, but the active discovery of opportunities for improving one's quality of life.

Searches for affordable housing in desirable communities could automatically generate "agents" to identify available properties; people with specific health issues might be notified of new treatment programs in their neighborhood; or people with an interest in Slovak-American culture could be notified of relevant lectures at the local community college.

The Semantic Web could be all the more powerful if it incorporates the "Internet of Things," the information generated by sensors and processors embedded in virtually all manufactured objects and connected with the World Wide Web. Cars, bridges, buildings, train tracks, sewer pipes, lawnmowers, winter coats, and Lake Michigan buoys might all benefit from two-way communication with interconnected systems. Without getting into too much detail, we can readily imagine how information can make almost all devices and all processes more efficient.

In such a world, community revitalization efforts would be supported by automatically updated information about physical infrastructure and aggregated statistics about health, education, criminal activity, and so on. Sure, this all depends on full participation by governments and institutions in a standardized system, and it depends on foolproof, unbreakable identity management. But after the awe-inspiring growth of web technologies since 1990, does anyone still doubt that such developments are possible over the coming years?

## Conclusion

The information needs of community intermediaries and their constituents are unlikely to be fully met by software companies, who prefer more lucrative markets. To some extent, mission-driven organizations need to adopt a do-it-yourself ethic when it comes to converting raw data into high-value retail information tools. Fortunately, raw data are becoming more available every day, and technologies for converting raw data to retail data products are becoming simpler and cheaper. But the will to create or invest in tools must be strong, the resources must be found, and relationships with civic hackers must be nurtured, or we will continue to operate in an information-poor environment. And crucially, a commitment to leveraging information tools, once they are created, must be embraced by intermediaries and individuals within the community. Information in itself has no value, but it can present opportunities for effective action. Clearly, acting on less information is easier than acting on the basis of data analysis. That should not come as a surprise to anyone who has adopted evidence-based decision processes. Good decision-making requires time, effort, and skill. But the better our tools become, the easier it will be to adopt data-driven action plans.

## REFERENCES

Executive Office of the President. Jan. 21, 2009. *President Obama signed the Memorandum on Transparency and Open Government.* Federal Register [74 FR 4685] http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment

Fitzek, F.H.P., T. Mikkonen, and T. Torp, eds. 2010. *QT for Symbian.* Hoboken, NJ: John Wiley and Sons.

Open Government Working Group. 2007. *8 Principles of Open Government Data.* https://public.resource.org/8_principles.html. Accessed September 1, 2014.

Sanders, Greg. The Business Case for Real-Time Sharing of Government Data. Presentation at the Urban and Regional Information Systems Association (URISA) conference. August 20, 2007. http://downloads2.esri.com/campus/uploads/library/pdfs/132035.pdf

Smith, Michael. 2008. *Building an Interoperable Human Services System: How Allegheny County Transformed Systems, Services and Outcomes for Vulnerable Children and Families.* Smithtown, NY: Stewards of Change.

Windley, Phillip J. 2003. A Web Services Manifesto (PowerPoint presentation). Accessed September 1, 2014. http://www.windley.com/docs/EnablingWebServices/Web%20Services%20-%20Jan%202003.pdf

World Wide Web Consortium. n.d. W3C Semantic Web Activity. Accessed September 1, 2014. http://www.w3.org/2001/sw/

# 4

# A Framework for Indicators and Decisionmaking

This chapter focuses on how community information is put to use for collective purposes. The payoff from work with community information derives from how it is applied and the influence its applications have on community change and societal goals. As noted in chapter 2, considerable benefit can be gained simply by releasing some types of data so that individuals can use them directly. For example, many local transit authorities make real-time bus schedule data available over the Internet. Yet more valuable are the 311 systems that let local governments track and respond to citizen requests for information and services with great efficiency (see Fleming's essay at the end of chapter 2).

However, this book focuses on applications designed to achieve broader collective ends at the local level; applications intended for use by institutions. A number of these applications are reviewed in chapters 5 and 6. With the expansion of community information and the advent of local data intermediaries, applications as illustrated by these examples have grown in their breadth and sophistication. They reflect the increasingly complicated world of governance at the local level.

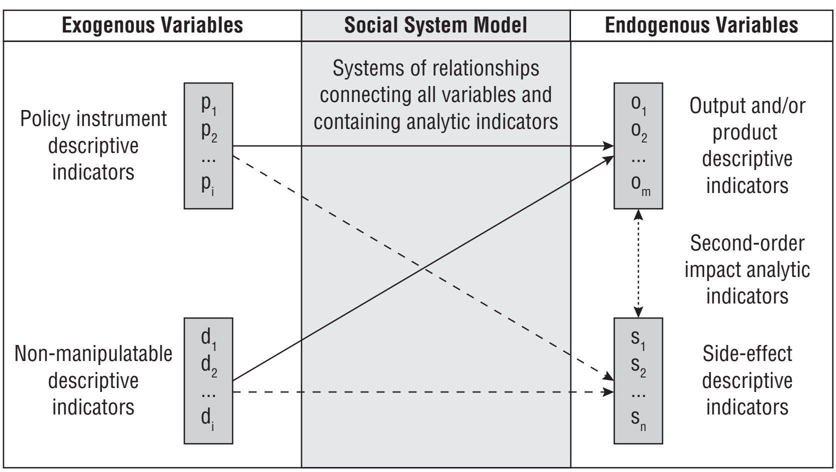
This chapter offers a framework for understanding these examples and others. We begin by reviewing the multiple uses of indicators presented in one of the foundational conceptualizations of the social indicators field. We then describe how this model translates into five basic types of applications that illustrate the scope of the community indicators work today.

## Social Indicators and Their Uses

Many think of the social indicators field as being mostly about selecting a set of outcome indicators (measures of the ultimate well-being of a community) and then monitoring trends in those indicators to see whether conditions are getting better or worse. However, this function (which we will call *situation analysis*) is only one of several that indicators can be used to perform. It has been recognized from the outset that a broader range of indicators (outcome indicators and other types) needs to be used in different ways and in other types of activities if sound public programs and policies are to be developed.

One of the early scholars in the field who emphasized this broader view was Kenneth Land (1975).[1] Land's conception identifies five types of indicators (see figure 4.1). The two types of indicators on the left of the diagram cause change in the social system: (1) policy instrument descriptive indicators (e.g., the magnitude of investment in, and other characteristics of, a workforce investment program) and (2) nonmanipulative descriptive indicators (also called context indicators) (e.g., the health of the metropolitan economy and other factors that an initiative cannot alter itself, but that will influence the success of a program).

**Figure 4.1.** Indicator Relationships in a Social System Model




*Source:* Land 1975.

The two types of indicators on the right side of figure 4.1 reflect results: (3) output or end-product descriptive indicators (the traditional outcome indicators that analysts or activists care about, such as more neighborhood residents being employed) and (4) side-effect descriptive indicators (not the key results being sought, but relevant things that occur because of the way the system operates in this context). In the middle of the diagram are (5) analytic indicators that speak to the cause-and-effect relationships between all the other types of indicators in action. As shown in the diagram, Land refers to this as the social systems model. In today's community initiatives, this is often referred to as the theory of change or, more simply, the logic model. This recognition of the varying types and functions of indicators is the basis for our framing of the main types of applications that are the foundations of work with community indicators (presented in chapter 2).

## Applications in the Community Indicators Field Today

We suggest that there are five basic types of applications that sometimes individually, or more often in combination with each other, underlie virtually all collective uses of community data. We briefly explain how each of them works and discuss forces now influencing how they are being conducted. Our conceptualization of the five basic types of applications is derived mostly from framing developed by Land (1975) as discussed above.

The uses of community information can be viewed as having systematic relationships to one another. Although a group or organization may well start with situation analysis, achieving impact will then require planning policies and programs to respond to the findings, implementing and monitoring those courses of action, and then researching cause-and-effect relationships to improve the model that serves as the basis for planning. Our five types of applications, then, are as follows:

- *Situation analysis* examines indicator values on conditions and trends to determine whether the circumstances in a community are getting better or worse, to identify problems and opportunities, and to gauge the relative importance of each. Land's model points out the importance of looking at context indicators as well as outcome indicators to help understand the forces behind observed changes.

- *Policy analysis and planning* use the data to formulate courses of action (adopting new laws as well as designing new programs) in response to findings from the situation analysis and to assess the advantages and disadvantages of alternatives. The process entails alternative courses of actions (represented by the policy instrument indicators, taking into account nonmanipulative descriptive indicators) and assessing their likely results (outcome and side-effect indicators) via a theory of change based on explicit assumptions about cause-and-effect relationships (analytic indicators).
- *Performance management and evaluation* assess trends in performance and relevant aspects of the context and use the data as a basis for making midcourse corrections and for subsequent evaluation. This means monitoring change for all the types of indicators in the system, reporting results, and in effect reentering an analytically based planning mode to adjust the course of action for the period ahead.
- *Education and engagement* use the information and analysis already developed in the categories above to influence (most often, to build support for) a proposed agenda. For example, the presentation of data may be used to engage residents in crime-prevention programs or make the case to government agencies, philanthropies, and the public at large for financial and other support for an agenda.
- *Neighborhood research* uses data to improve the understanding of how the overall social system works by examining changes in indicators and their interactions. This sheds light on the patterns of cause and effect through which social interventions and other forces interact to influence outcomes. One of the key forces is the effect of neighborhood conditions on individual outcomes.

In today's best practice, the process of information use in managing a collective agenda is quite consistent with Land's concept. The basic elements of this process work whether the agenda at hand is the limited work program of one social service agency or that of a large-scale, long-term, multifaceted, collaborative effort to improve distressed neighborhoods.

Below we offer descriptions of each of the five types of applications to clarify how they work, and we also note suggestions being made today to improve practice in each area. Chapters 5 and 6 will provide more in-depth examples of the first four categories, and chapter 7 will discuss analytic issues and methods related to neighborhood research.

## Situation Analysis

Valuable examples of situation analysis can be quite simple. For example, a local data intermediary releases information to the public on one or two new facts that are surprising and inherently suggest a need for response—perhaps that "three neighborhoods on the lower west side saw an unexpected jump in the vacancy rate (or crime rate) that did not occur in other neighborhoods in that district."

In more comprehensive examples, indicators are selected from all topical domains regarded as important to the community's well-being, and the data are recurrently updated and reviewed. This is the approach most often aspired to by community indicators projects, as discussed in chapter 6. The purpose of these projects is to give the community an accurate sense of whether the trends are positive in each topic area and, thereby, to help the community establish priorities for response.

As to improving practice in situation analysis, the most important theme revolves around analysis that will help users set priorities and then translate the results into action. A valuable framework for moving in that direction is the practice borrowed from strategic planning: SWOT analysis. It is called that because the user group must explicitly assess its own *s*trengths and *w*eaknesses in relation to both the *o*pportunities and *t*hreats the environment appears to hold. Doing the latter means that context indicators must be reviewed as well as outcome indicators. For example, regardless of recent trends in neighborhood housing values (outcome indicators), it is important to note recent trends in the metropolitan housing market (context indicators) to realistically understand potentials for neighborhood values in the future.

## Policy Analysis and Planning

Policy analysis and planning involve using community information to design the actions to be taken in response to the findings of the situation analysis. Consistent with the concept of strategic planning [see, for example, McNamara (2009)], this process begins with a formal review of policy or program options to find a set that seems likely to work best in light of what has been learned and priorities that have been identified. In other words, what courses of action might we follow to respond to threats and opportunities in light of a better understanding of our strengths and weaknesses?

This process virtually always entails some type of mental testing of alternative scenarios, and even the best of these work by trial and error. The process may start with clarifying the theory of change or logic model (the social systems model in Land's diagram)—that is, clarifying and revising ideas about how the relevant patterns of cause and effect actually work. After this clarification, alternative sets of promising interventions are formulated and, consistent with the theory, expectations about the advantages and disadvantages of each alternative are made explicit. Thinking through the implications may lead to recognizing new problems and opportunities, which may then suggest directions for adjustments to the first plan (a new scenario).

In local policy and community work, the pressure from funders and leaders in the field for more data-driven decisionmaking in this process (weighing the options and deciding what actions to take) is now substantial. This pressure implies increasing efforts to take advantage of community information to quantify expected implications of alternative scenarios. Instead of saying in general that one option is likely to be less risky, less expensive, or more effective than another, the team tries to use model relationships to project trends under varying assumptions and to estimate costs. Given the complexity of local socioeconomic systems, no one has yet come close to developing measures for a system as a whole, but it is now reasonable to try to measure much more than in the past. These efforts will be wise to employ the framework of cost–benefit analysis [see, for example, Boardman et al. (2001)], even if all the desired parameters cannot be estimated reliably.

The practice of formulating and testing alternative scenarios has been a recommended part of strategic planning for half a century.[2] The practice has evolved slowly because the work is complex and there has been an unfulfilled expectation that the art of predicting the future would improve enough so institutions would not have to think through very many alternatives. Now there seems to be more acceptance that we live in an uncertain world. The future of complex systems (like cities and neighborhoods) will remain very difficult to predict reliably. However, computer-based tools are being developed to simplify the task of scenario testing, so more work of this type seems probable (Avin 2012). The task is no longer one of preparing a single plan that one expects to hold constant, but rather developing the capacity to figure out how to change the plan if B happens or C happens, instead of A as was first assumed.

## Performance Management and Evaluation

The work defined above yields a plan (a particular scenario) that is selected to guide operations at least initially. The next stage in the process, which is also data intensive, is monitoring performance against that plan. This first means collecting and displaying updated data on the program's outcome indicators that tell directly whether the key results are being achieved.

In recent years there has been substantially increased emphasis on "results-oriented" behavior in the social sector and performance measurement to back it up. Those advocating this direction recognize that it may require a major culture change for many social service agencies, but they see much more quantification along these lines as essential if underlying goals are to be met (Morino 2011). And new approaches to measurement and assessment have been developed to help practitioners assess performance more realistically; see, for example, results-based accountability as proposed by Friedman (2005).

But, again, simply looking at changes in outcome measures is not enough. It is also necessary at the same time to examine measures that describe the actions that have been taken, which may or may not be fully consistent with what the plan specified, and changes to the context indicators (some exogenous event or trend may have deflected performance, despite good work by the initiative team). In this phase, managers are thinking through why the outcomes turned out as they did.

Ideally, the meeting in which the performance review takes place should not just be about looking at the data and handing out praise or blame. It should also be purposefully designed to include serious thought about what worked, what did not, and why (explicitly referencing and possibly amending the theory of change as appropriate). They can then make midcourse corrections to the plan, or possibly revise it dramatically, on the basis of that analysis. In other words, additional scenarios should be formulated and tested, spurred on by the information updates.

But this type of performance review does not happen often enough. In fact, public-sector agenda management in America has been criticized of late because (1) strategic plans have become rigid and too isolated from the processes of regular management, and (2) performance measurement has become an end in itself, not generating enough of the innovative thought that goes into devising more effective courses of action. Poister (2010, 246) argues that

> making strategy more meaningful in the future will require transitioning from strategic planning to the broader process of strategic management, which involves managing an agency's overall strategic agenda on an ongoing rather than an episodic basis, as well as ensuring that strategies are implemented effectively.... We need to shift the emphasis of the performance movement from a principal concern with measurements, to the more encompassing process of performance management.

### *PerformanceStat*

A trend that is consistent with Poister's suggestions is the PerformanceStat movement that has taken hold in a number of cities and higher levels of government around the country. As described by Fleming in her essay at the end of chapter 2, this movement first emerged in full multidepartment form as CitiStat in Baltimore.[3] It entails the preparation of considerable data on the performance of individual departments, and then intensive regular review sessions with high-level officials. In addition to reviewing whether performance targets have been met, it also emphasizes devising and proposing solutions when service problems are detected. *PerformanceStat* is the global term employed by Behn (2008) to describe these types of programs.

Most often, PerformanceStat programs rely on data generated by the department at hand. Even looking at the department's output data, recognizing neighborhood variations is important (e.g., the performance of road maintenance departments in distressed communities has often been below the standards for the rest of the city). But goals and performance targets should relate not only to outputs (the amount of a service produced), but also to impacts (results). Understanding changes in impacts almost always implies the need to look at context indicators and a broadened set of outcome indicators as well, and that is likely to expand demand for community information beyond the data holdings of individual departments. And, again, knowledge of neighborhood variations is likely to be critical in assessing the results.

## *Collective Impact*

There has long been a concern that attempting to address fundamental problems of American society through a collection of fragmented individual programs is not going to work. But the conceptualization of a forceful alternative has emerged only recently.

To be sure, performance management can improve individual programs. But other goals are broader than any one level or branch of government (let alone any one program) can handle. These agendas require cross-sector engagement and support; that is, they need to be taken on by civic leadership as a whole (i.e., by governing coalitions, as we have discussed them in chapter 2).

This broader approach is called *collective impact.* Its most prominent example so far is the Strive partnership, a Cincinnati, Ohio, coalition involving child care advocates, school district superintendents, college presidents, local philanthropies, business leaders, and social service providers (Bridgespan Group 2012). Instead of focusing on the stated objectives of their individual organizations, they have pledged to collectively pursue, and be held accountable for, achieving broader educational goals. Their mission is to "coordinate improvements at every stage of a young person's life from 'cradle to career'" (Kania and Kramer 2011, 36).

One of the required conditions of the collective impact approach is shared measurement.[16] "Collecting data and measuring results consistently across all participants ensures efforts remain aligned and participants hold each other accountable" (Hanleybrown, Kania, and Kramer 2012, 1). The collective impact approach is spreading beyond the realm of education. For example, it has been recommended to address citywide community development needs in a proposal by Brophy (2011) for Milwaukee, Wisconsin. As more local coalitions adopt the collective impact approach, that will also expand demand for a wider array of reliable and frequently updated community information.

## *Evaluation*

Where performance management emphasizes finding out how well a program is working in the short term as a basis for making midcourse corrections, program evaluation seeks to find out whether the program was successful in meeting its objectives related to effectiveness and impact over the longer term (Walker and Moore 2011; Auspos and Kubisch 2012). There is a substantial literature on evaluating local social programs and community improvement initiatives; see the discussion in chapter 7 and, for example, Harrell et al. (1996) and Connel et al. (1995).

Two trends in the field of evaluation should increase the demand for community information in the years ahead. First, although in the 1990s funders focused much more on evaluation, they now give more emphasis to performance management in the local initiatives they fund. Walker and Moore (2011) note that whereas evaluation is often performed by external researchers for an outside audience after the program is complete, performance management is conducted by the managers and staff of the initiative itself. It is "owned by" them and conducted to achieve their key internal objective of improving performance in process. In addition to increasing demand for community information in the short term, systematic data collection from the start could lead to the availability of better information for evaluation in the longer term.

Second, alternative approaches to evaluation are being considered. Evaluators would like to find out whether a program unambiguously caused the relevant outcomes that are observed, and a full randomized control trial is the ideal way to accomplish that. Yet controlled trials for complex community initiatives are extremely difficult to implement. Chapter 7 describes a number of alternatives that, in today's data-rich environment, are less costly to implement than they were in the past. In addition to creating more demand for neighborhood-level data, exploring new analytic options should improve the methods for using them.

## Education and Engagement

One of the central determinants of the actions we take, individually and collectively, is our perception of how the world works and how conditions are changing around us. Facts that run counter to the conventional wisdom are likely to command attention and, ultimately, change behavior. Community information generated in all the types of applications above can influence public opinion and motivate action.

Organizations can adopt a range of tactics to educate the public at large, including the traditional media, blogs, and public events. Other advocacy efforts target specific audiences. For example, an activist may use neighborhood analysis about increased housing prices in her testimony to the city council about potential city programs to promote homeownership.

Neighborhood indicators can also be a tool in more active engagement of stakeholders. As noted above, local data intermediaries often use community information to build collaboration and support. First, data and analyses are developed in the situation analysis stage, and then, new proposals with facts to back them up result from the planning stage. But then, the proposals and facts are presented to external audiences to encourage their participation and support, making a case that is now sounder and more compelling than it would have been without the analysis.

In other cases, it is not the finished analysis that has been the key, but the process of developing the data. For example, in a number of instances organizers have involved neighborhood youth in conducting surveys of conditions in their communities. What the youth learn from their neighbors in this way binds them into participation in the subsequent initiative mounted to improve the conditions they have studied.

Interviews with many local NNIP clients point to how important the data themselves are to the success of these efforts. The Strive initiative in Cincinnati, noted earlier, illustrates the role of data in getting key partners to the table and keeping them there. The Bridgespan Group (2012, 26) states

> [I]n its work to date, Strive has found it hard to engage either residents or its regional partners around the goal of building a civic infrastructure, which comes across as theoretic and insufficiently tangible. Engaging them around specific outcomes, such as graduation rates, has been extremely successful, however. In these instances, data has provided the focal point and lingua franca that binds the interested parties together. In contrast, Strive leaders report that when they briefly lost their focus on data, the momentum and progress of the partnership lagged.

## Neighborhood Research

Few would disagree that to develop effective strategies for improving community conditions, policymakers and planners need to improve their knowledge of the process of neighborhood change. Land's model emphasizes that outcomes are determined not only by interventions, but by interventions interacting with other forces (nonmanipulative descriptive indicators) via processes that themselves may vary across communities. Some theory of change in this context is essential to decisionmaking. The more complete and reliable that theory becomes, the better the chances should be for successful interventions.

In addition to changes in neighborhoods, the field needs to understand the changing circumstances of people. A growing body of evidence demonstrates the effects of neighborhood conditions on individuals and families, as discussed in more depth in chapter 7. Researchers must use complex analytic techniques to sort out the effects of neighborhood conditions from those of individual and family characteristics and behavior. Better knowledge of the mechanisms by which neighborhood conditions help or harm people's life chances can help policymakers and practitioners prioritize interventions to improve the most influential conditions. Research on neighborhood change and effects has advanced markedly over the past two decades, but methodological hurdles remain. Chapter 7 reviews selected techniques that represent promising approaches to a deeper understanding of both processes and explores the state of the art of neighborhood indicators analysis for four purposes: description, classification, prediction, and explanation.

## Conclusion

Using indicators to monitor community well-being (part of situation analysis) is important. But this chapter has shown that several other types of applications are critical as well. A considerably more robust view of the potential uses and usefulness of indicators was put forward by Land when the social indicators movement first got under way, and the features of that approach are the seeds of the types of applications that are emerging as essential for data work in local governance today.

Many clients for this work—ranging from individual neighborhood groups to civic leaders—are now pushing for smarter designs and accountability in local initiatives and programs, and they are increasingly recognizing that data-driven decisionmaking may be key to achieving their goals. This emphasis means that the years ahead are likely to see substantially expanded demand for the effective production and use of community information.

The decision-oriented applications are likely to be emphasized: sharper diagnosis of community problems and opportunities (situation analysis), more effective design of action programs in response (policy analysis and planning), and insistent monitoring and assessment of results (performance management). But the other applications we have noted will be essential as well: education and engagement to assure the data get used, and evaluation and research to provide more reliable understanding of how neighborhoods function and how community interventions work to affect results.

How well is the field moving toward this vision? The next two chapters provide examples that offer a basis for making that assessment.

## NOTES

1. The need for this broader view is also reflected in lesson 8 of the review of the social indicators field by Cobb and Rixford (1998, 23).

2. See a discussion of its recommended use in national security planning by Brown (1968). Also see Quay (2012).

3. New York City's CompStat, which focused only on police department performance, was actually the initial model (Behn 2008).

## REFERENCES


Auspos, Patricia, and Anne C. Kubisch. 2012. *Performance Management in Complex, Place-Based Work.* Washington, DC: Aspen Roundtable on Community Change.

Avin, Uri. 2012. "Tools for Building Scenarios: Sorting Out What to Use When." *Planning* December: 40–43.

Behn, Robert D. 2008. *The Seven Big Errors of PerformanceStat.* Rappaport Institute/Taubman Center policy brief. Cambridge, MA: John F. Kennedy School of Government, Harvard University.

Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer. 2001. *Cost Benefit Analysis: Concepts and Practice.* 2nd ed. Upper Saddle River, NJ: Prentice Hall, Inc.

Bridgespan Group. 2012. *The Next Generation of Community Revitalization: A Work in Progress.* Boston: The Bridgespan Group, Inc.

Brophy, Paul C. 2011. *Community Development in Milwaukee: An Assessment and Recommendations.* Baltimore: Brophy and Reilly, LLC.

Brown, Seyom. 1968. "Scenarios in Systems Analysis." In *Systems Analysis and Policy Planning: Applications in Defense,* edited by E. S. Quade and W. I. Boucher. Report R-439-PR (abridged). Santa Monica, CA: Rand Corporation.

Cobb, Clifford W., and Craig Rixford. 1998. *Lessons Learned from the History of Social Indicators.* San Francisco: Redefining Progress.

Connel, James P., Anne C. Kubisch, Lisbeth Schorr, and Carol H. Weiss, eds. 1995. *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Context.* Washington, DC: Aspen Institute.

Culhane, Dennis P., John Fantuzzo, Heather L. Rouse, Vicky Tam, and Jonathan Lukens. 2010. "Connecting the Dots: The Promise of Integrated Data Systems for Policy Analysis and Systems Reform." *Intelligence for Social Policy* 1 (3): 1–9.

Friedman, Mark. 2005. *Trying Hard Is Not Good Enough: How to Produce Measurable Improvements for Customers and Communities.* Victoria, BC.: Trafford Publishing.

Hanleybrown, Fay, John Kania, and Mark Kramer. 2012. "Channeling Change: Making Collective Impact Work." *Stanford Social Innovation Review* Winter.

Harrell, Adele V., Martha R. Burt, Harry P. Hatry, Shelli B. Rossman, and Jeffrey A. Roth. 1996. *Evaluation Strategies for Human Services Programs: A Guide for Policymakers and Providers.* Washington, DC: Urban Institute.

Kania, John, and Mark Kramer. 2011. "Collective Impact." *Stanford Social Innovation Review* Winter.

Land, Kenneth C. 1975. "Social Indicator Models: An Overview." In *Social Indicator Models,* edited by Kenneth C. Land and Seymour Spilerman. New York: Russell Sage Foundation.

McNamara, Carter. 2009. *Strategic Planning in Nonprofit or For-Profit Organizations.* Managementhelp.org/plan-Dec. accessed October 27, 2009.

Morino, Mario. 2011. *Leap of Reason: Managing to Outcomes in an Era of Scarcity.* Washington, DC: Venture Philanthropy Partners.

Poister, Theodore. 2010. "The Future of Strategic Planning in the Public Sector: Linking Strategic Management and Performance." *Public Administration Review* 70 (December): s246–s254.

Quay, Ray. 2012. "Brave New World of Scenario Planning." *Planning* (November): 52.

Schorr, Lisbeth B., and Frank Farrow. 2011. *Expanding the Evidence Universe: Doing Better by Knowing More.* Washington, DC: Center for the Study of Social Policy.

Walker, Karen E., and Kristin Anderson Moore. 2011. "Performance Management and Evaluation: What's the Difference?" Research to Results brief #2011-02. Washington, DC: Child Trends.

# 5
# Using Data for Neighborhood Improvement

Data on conditions and trends for small areas (neighborhoods) are essential for effective policymaking in cities and metropolitan areas. This information can be used in many ways, but an important distinction for the purposes of this and the next chapter is whether the focus of the work is within neighborhoods or whether the data are being applied across neighborhoods. Applications supporting decisions across neighborhoods are examined in chapter 6. In this chapter, we focus on experience in using neighborhood-level data to improve conditions within individual neighborhoods.

Chapter 1 points out that adverse conditions in low-income neighborhoods motivated the establishment of the original local data intermediaries in the early 1990s. This chapter describes how data have been used in attempts to revitalize such neighborhoods and discusses lessons learned from these experiences. Before we examine these specific examples, we review the evolution of programmatic approaches to neighborhood improvement and the role that data play in these approaches.

## Community Improvement: Evolution of Concepts and Practice

Although efforts to revitalize poor urban neighborhoods had been initiated by governments before, the seeds of probably the most notable expansion in scope resulted from the work of Robert F. Kennedy and others in the mid-1960s. They married the idea of developing indigenous neighborhood-controlled entities as the change agents [nonprofit community development corporations (CDCs)] with a major effort to tap the wealth of the nation's major private corporations to support their work. The first CDC, the Bedford Stuyvesant Restoration Corporation, was established in Brooklyn, New York, in 1966.

Another innovation was essential for expansion: the creation of independent intermediary organizations that would help raise the needed private capital and provide technical assistance as well as loans and grants to the local organizations. The largest of these, the Local Initiatives Support Corporation (LISC), was started with grants from the Ford Foundation and several large corporations. Others include the organizations now known as NeighborWorks and Enterprise Community Partners.[1] The intermediaries leveraged their resources, drawing in additional private funding from local sources, and the federal government provided more support in various forms.[2] Since the 1970s, the number of CDCs expanded markedly in most large US cities, reaching 4,600 in 2005 (National Congress for Community Economic Development 1989, 2005), as did their production capacity and accomplishments (Walker 2002).

Through the mid 1980s, community development generally focused on improving the housing stock and commercial activity in neighborhoods through new construction or rehabilitation of individual properties. At about that time, however, a number of practitioners advocated for and began to implement a more comprehensive form of neighborhood improvement. First, community-building efforts, mostly funded by private foundations, gave more emphasis to resident engagement, building neighborhood institutional capacity and human development programs (Kubisch et al. 1997, 2002; Kingsley, McNeely, and Gibson 1997). In this approach, building indigenous community capacity to manage neighborhood improvement was given high priority.

Eventually, some of the organizations that were part of the traditional community development movement expanded their horizons, merging the bricks-and-mortar side with resident engagement and programs to address education, health, public safety, and other community needs.[3] The more comprehensive approach was motivated by the recognition that the problems of troubled neighborhoods are multidimensional and interrelated and that work in any one field alone, such as housing improvement, would clearly not be enough to be transformative. Kubisch et al. (2010) have identified 48 major community change efforts initiated since the early 1980s.

To be sure, the potential of neighborhood improvement initiatives has been questioned over the years. Although the exact path for the future is not clear, several reasons suggest that, even in a resource-constrained environment, support for these approaches will continue in some form. First, the most important criticisms of the past now seem to have been muted. One example is the argument made by some (starting in the 1980s) that poverty reduction should focus on people-based programs instead of place-based programs. It is now difficult to find policy discussions in the literature that do not recognize the need for both (Duke 2012). Another example is the argument that the emphasis should be on mobility strategies that, instead of focusing on trying to fix troubled neighborhoods, facilitate the movement of the poor out of them and into "neighborhoods of opportunity" elsewhere. Here too, however, these seemingly contrasting approaches are now more often being seen as complementary. In the view of Pastor and Turner (2010, 1), "we need a broader portfolio of 'place-conscious' strategies that simultaneously improve neighborhood conditions, open up access to opportunity-rich communities, and realign regional growth and development to better connect low-income people and places with regional opportunities."

A second boost for neighborhood improvement has been the recent engagement of the public health sector. The field's new emphasis on prevention has been coupled with a growing understanding that preventing some of the nation's most intractable health problems requires addressing the problems of distressed neighborhoods. Marjorie Paloma of the Robert Wood Johnson Foundation points out, "This growing recognition of housing, neighborhoods and factors such as income and education—the social determinants of health—has led the health sector, and increasingly the housing and community development sectors, to look beyond improving access to health care to address root causes to help people avoid getting sick in the first place" (Paloma 2012).

Finally, the Obama administration has raised the profile of place-based approaches through prioritizing policies supporting neighborhood improvement. Office of Management and Budget Director Peter Orzag called on all federal agencies to explicitly examine and enhance the "place" impacts of their programs (Orzag et al. 2009). Furthermore, as noted in chapter 2, the Obama administration launched several new initiatives that call for cross-agency work within low-income neighborhoods, including Promise Neighborhoods (Tough 2008; Promise Neighborhoods Institute 2011; Biglan et al. 2011; US Department of Education 2012), Choice Neighborhoods (US Department of Housing and Urban Development 2011a, 2011b), and the umbrella Neighborhood Revitalization Initiative (White House 2011).

## The Role of Data in Strengthening Communities: A Review

As noted in chapter 1, Ahlbrandt and Brophy called for better data to guide this growing movement in 1975, but the response was initially negligible. One reason may have been that multitopic neighborhood-level data did not seem essential to the bricks-and-mortar, project-by-project world of community development at that time. Probably more important, however, were the severe constraints that still remained on the supply side. The technical and institutional innovations that supported producing and sharing administrative data and building geographic information system (GIS) capacity were not present until the late 1980s.

However, once such systems started to be built, they were used productively, if haltingly, in various ways even in the earliest comprehensive initiatives. Over the years, both the frequency and sophistication of data use have increased in community work. Experience shows examples of all the basic types of applications we introduced in chapter 4: situation analysis, policy analysis and planning, performance management, education and engagement, and understanding neighborhood effects and dynamics. Most often, projects have involved two or more of these applications.

Still, even today the community revitalization field cannot be characterized as data driven overall. Most of the practitioners and neighborhood residents who steer these initiatives are not yet regularly using data in these types of applications. The recent examples discussed below suggest progress in using data, but also demonstrate the barriers still to be overcome.

The review of experience in the remainder of this chapter discusses uses of data in recent neighborhood improvement efforts that illustrate themes that the authors judge to be key elements of what this field has accomplished and where it is headed. The review is organized into six sections, sequenced roughly according to stage of development:

- *Efforts in community initiatives to assess neighborhood conditions and trends.* Community initiatives began to use neighborhood indicators systems in an exploratory mode soon after they were developed and also collected original data on conditions not easily captured by administrative data.
- *Uses of data to learn and mobilize around key issues.* Some influential uses have focused on individual topics, using data to better understand an issue and then to mobilize external support for neighborhood recommendations.
- *Sophisticated uses of property data in the foreclosure crisis.* Data systems with a wide array of information about properties have enabled more nuanced decisions about how to address the problems of neighborhoods hit hard by foreclosure and abandonment.
- *Applications in large-scale comprehensive initiatives: examining the potential for performance management.* The largest comprehensive community development program operating in any US city in the 2000s was Chicago's New Communities Program (NCP), notable for its pioneering efforts to apply performance management in the community development context.
- *New insights on neighborhood change from Making Connections.* The Annie E. Casey Foundation's Making Connections initiative focused on improving low-income neighborhoods in 10 cities and supported an associated survey research program that has yielded important new insights related to neighborhood change.
- *Intensive data use in program planning and implementation: Promise Neighborhoods.* The Promise Neighborhoods Initiative has upgraded expectations for the use of data in program planning and implementation and suggests future directions for the field.

## Efforts in Community Initiatives to Assess Neighborhood Conditions and Trends

Chapter 1 noted that the first ongoing multisource neighborhood-level data system was developed by a community-oriented research center at Case Western Reserve University in Cleveland, Ohio. The system was established in part to support the new Cleveland Community Building

Initiative (CCBI), which aimed at revitalizing four of that city's most troubled neighborhoods consistent with community-building principles (Cleveland Foundation Commission on Poverty 1992). The system was developed around a set of indicators selected in the context of CCBI that covered a wide range of topics under five broad headings: economic opportunity; institutions and services; family, child, and youth development; safety and security; and neighborhood identity and pride. Most of the data came from the new flow of local administrative files, although some were derived from the Decennial Census and special surveys (Coulton 1995).

There have been no formal studies of how the data were used in CCBI, but interviews with center staff who were involved clarify the basics.[4] Data from the system were presented in charts and tables that showed the indicator values for the four neighborhoods (called villages in CCBI) individually and for the city as a whole. The data were first shown to practitioners who served as village coordinators. In interactive sessions, they examined where their village stood in relation to the other villages and the city as a whole on indicators of greatest interest. Learning where they were comparatively better or worse off yielded ideas as to where their programmatic priorities ought to be set. The coordinators would often note results that surprised them, then ask center staff to go back and get more data to explain or amplify the findings, leading to additional exploration of possible causes and implications.

In some cases, coordinators brought the stories back to neighborhood residents to stimulate their thinking about directions for plans and projects. One of the particular concerns at the time was juvenile delinquency and questions about better ways to keep teens occupied after school. System indicators (e.g., crime incidence, school attendance rates) showed the basic patterns and trends. Follow-up observation by residents identified when and where youth tended to loiter, facilitating outreach by program providers.

The other original NNIP partners (Atlanta, Georgia; Denver, Colorado; Oakland, California; and Providence, Rhode Island) reported similar collaborative approaches as their own indicators systems were being developed (Kingsley 1999). Rather than simply presenting findings, they all used the data as a way to start conversations. After introducing the statistics, they would ask residents and initiative managers questions such as, "Does this finding sound realistic for your neighborhood?" "What other information should we be looking at?" "What do you think are the forces at work that make it turn out this way?" and "What do you think you could do about it?" This approach is the centerpiece of the NNIP "democratizing information" idea as introduced in chapter 2. Again, the objective of this process is that the residents and managers feel they own the data and the conclusions drawn from them.

This approach to the use of data can be seen not just as a way to select programmatic themes but as a central part of the intervention in and of itself. In much community work, an overarching goal has been to build community capacity (Chaskin et al. 2001; Kubisch et al. 1997; Gittel and Vidal 1998). Neighborhood residents and practitioners develop new skills as they question the meaning of the data and think through implications for action. They build their own capacity to design and manage revitalization efforts more effectively.[5]

Participants in other community initiatives were also experimenting with data to further their objectives in various ways. Chaskin's essay at the end of this chapter reviews the difficulties as well as the achievements of several of them that integrated data use into their programs. He discusses how community groups participating in the Ford Foundation's four-site Neighborhood and Family Initiative all took different approaches when required by the sponsor to conduct the local part of the initiative's evaluation.

Chaskin also describes a collaboration between community organizations in two Chicago neighborhoods and a university partner that entailed surveys to measure aspects of social capital. This work involved all parties jointly questioning the data and their implications in a manner similar to the NNIP examples noted above.

One of the Neighborhood and Family Initiative sites in Chaskin's essay undertook asset mapping, an activity that gained popularity beginning in the 1990s. Recognizing that most work with distressed neighborhoods was oriented around problems, asset mapping [developed by Kretzmann and McKnight (1993)] focused instead on the positives. It made the case that improvement was more likely if initiatives started with an inventory of the neighborhood's generally underappreciated assets. Such assets included, among others, voluntary associations, arts and culture, local businesses, resident skills, and valued buildings and spaces. Indicators from administrative data contributed to identifying location-specific assets, but identification required other methods of data collection as well, such as interviews and observational surveys. Considerable guidance has been offered on how to conduct asset mapping; see, for example, Dorfman (1998), Green and Haines (2002), and many of the publications of Northwestern University's Asset-Based Community Development Institute.[6]

## Uses of Data to Learn and Mobilize around Key Issues

We have mentioned the growing emphasis on a comprehensive approach in neighborhood improvement initiatives. Comprehensiveness, however, does not mean doing everything at once—a logistical impossibility. The only practical way to proceed is to begin working on a small number of priority opportunities and then add others strategically as earlier tasks are completed or phased down. At any point in time, activists and organizations are not working on all the topical areas they think are important, but eventually they will touch most of the bases. Priorities should generally be set based on an objective assessment of comparative importance across the full spectrum of problems and opportunities being faced by the neighborhood rather than being restricted to any one, or any small set of, narrowly defined programmatic specialties.

Some of the supporting data work that is needed will be comprehensive, such as recurrent reviews of trends in indicators of neighborhood well-being across all topics. But much of this work will be composed of data gathering and analysis supporting individual priority programmatic efforts. One group of examples in this category uses the data to assess patterns of demand or need within a neighborhood so managers can do a more sensible job of targeting locations for service delivery. In Milwaukee, Wisconsin, a youth service agency on the city's west side became aware that the number of female-headed households with children had grown rapidly in the area since the late 1980s. At their request in 1993, the local data intermediary (the Nonprofit Center of Milwaukee) mapped the density of children by block. The agency then used the maps both to justify the need for a new "First-Time Parents" counseling and service program and to target outreach efficiently (Kingsley et al. 1997, 66).

Another group in this category involves using data to better understand what is driving an important issue in a neighborhood and then highlighting the data in mobilizing support for action to address it. Among these, probably the most noteworthy impacts have been achieved when grassroots groups have motivated and taken charge of the work. In the mid 1990s, for example, a community organization in Camden, New Jersey, had been trying unsuccessfully to get the police to focus their attention around vacant buildings, where the organization's experience indicated that crime was most concentrated. At their request, a local professor prepared maps showing the overlay of vacant buildings and crime in the city. The maps indicated powerful correlations, and, when they were presented at a public meeting, the press picked up on the issue. This publicity stimulated further deliberations that ultimately led the police to revise their strategy in accordance with what the community group had suggested. The maps had clearly made the difference (Schmitt 1997).

In another case, the Neighbors Reaching Out Neighborhood Association in Nashville, Tennessee, took advantage of a local government offer to make recommendations on the allocation of Community Development Block Grant funds. Neighbors Reaching Out met with residents of their neighborhood to talk about priorities. Many of their suggestions focused on the need to improve housing with poor insulation and malfunctioning plumbing systems. Neighbors Reaching Out decided to recommend priority for home repair subsidies to senior citizens who owned homes that were poorly insulated and had older pipes that needed replacement. To be efficient in targeting outreach and project selection, they asked the Neighborhoods Resource Center (the local NNIP partner) to use data to identify properties likely to meet these criteria. Neighborhoods Resource Center staff used parcel data from the Davidson County Assessor's Office to identify neighborhood homes that, based on their age, were likely to contain deteriorated plumbing systems, settling on around 300 neighborhood properties that were built prior to World War II as a good approximation. They then merged age data from voting records to identify 87 of the 300 properties that were likely to be occupied by an individual over the age of 62. Neighbors Reaching Out found that this data analysis gave their requests substantial credibility in the funding allocation process (Cowan 2007).

A further example is the work of the Homewood Children's Village in Pittsburgh, Pennsylvania, in an initiative that began in the late 2000s and was modeled on the Harlem Children's Zone[7] approach. From conversations with residents, its leaders understood that physical deterioration of the Homewood neighborhood was viewed as a safety problem for its children. Many children had to walk past vacant and unsecured properties on their way to school. The University of Pittsburgh's School of Social Work was a key partner in the effort from the start and followed the principles of community-based participatory research, which they term "data-driven organizing" (Teixeira and Wallace 2013). The staff at the University Center for Social and Urban Research (the NNIP partner) was asked to help supply property maps and design a survey to reliably document housing conditions. The survey, fielded by neighborhood residents and university students, showed that the problems were not evenly spread. The group focused on the 30 worst properties (which became known as the "dirty thirty") and mobilized other residents to call the 311 citizen service line to advocate for action by the city on the targeted set. The campaign proved effective. Within about a month, 23 of the 30 had been demolished, boarded up, or improved.[8]

## Sophisticated Uses of Property Data in the Foreclosure Crisis

The devastating neighborhood effects of the foreclosure crisis of the late 2000s motivated more advanced examples of the use of data in addressing community issues.[9] These uses entailed assembling a broader array of relevant data on properties than is available in many local government information systems, and working collaboratively with a variety of stakeholders.

In places with weaker housing markets, the foreclosure process can result in long-term vacant housing, which then deteriorates due to poor maintenance. Research has shown that as vacancies increase in a neighborhood, so does the likelihood of crime and other social problems. The problems spread. As signs of disinvestment become evident in a few structures, property values erode, and more properties may be abandoned or otherwise become vacant. These problems have seriously negative impacts along a number of dimensions on the well-being of the families, who are either directly affected or living nearby.[10]

The foreclosure crisis came on quickly in 2007 and 2008, and most local governments were ill-prepared to deal with it. Some research documented the spatial patterns of foreclosure across cities and metropolitan areas, but other efforts used data to address the problem within neighborhoods. At the start, the priority at this level was mostly to gain some reliable understanding of the extent of the problem in a neighborhood and its effects on physical conditions overall. One example response, by the Southwest Organizing Project in Chicago, Illinois, is reviewed in Chaskin's essay.

A second example is a project in the Hickory Hill neighborhood of Memphis, Tennessee, which was undertaken by NNIP's Memphis partner, the Center for Community Building and Neighborhood Action at the University of Memphis. The center staff had access to administrative data on foreclosures, but they and neighborhood leaders recognized that information alone was unlikely to be enough to make a compelling case for intervention. Accordingly, community residents collected additional descriptive information on the characteristics of problem properties. These additional data enabled the community to present clearer evidence on how foreclosures were contributing to neighborhood blight. However, many blighted properties were not in foreclosure, so administrative foreclosure data were insufficient to prioritize city inspections. Hickory Hill leaders then used the additional data they had assembled, working with local government to better target limited city enforcement resources and to put pressure on owners of chronic problem properties to make improvements. The center was later asked to extend this work and prepare a citywide problem property audit (Buchanan et al. 2010).

In another example, the Hilltop Alliance, representing several neighborhoods in South Pittsburgh, Pennsylvania, asked the University Center for Social and Urban Research to conduct a fairly comprehensive analysis of the foreclosure problem in their area and to recommend both preventative and remedial actions. The research documented trends in foreclosure and property sales that were useful in considering responses. For example, it showed that 68 percent of foreclosures were filed by just 10 lenders, only one of which was based in Pittsburgh; also, that a surprisingly high share (38 percent) of properties involved in foreclosure sales had been purchased by investors (rather than intended owner–occupants). The University Center recommended that the Alliance promote housing counseling to homeowners and serve as a broker to transfer properties to responsible new owners or a land bank (University Center for Social and Urban Research 2011).

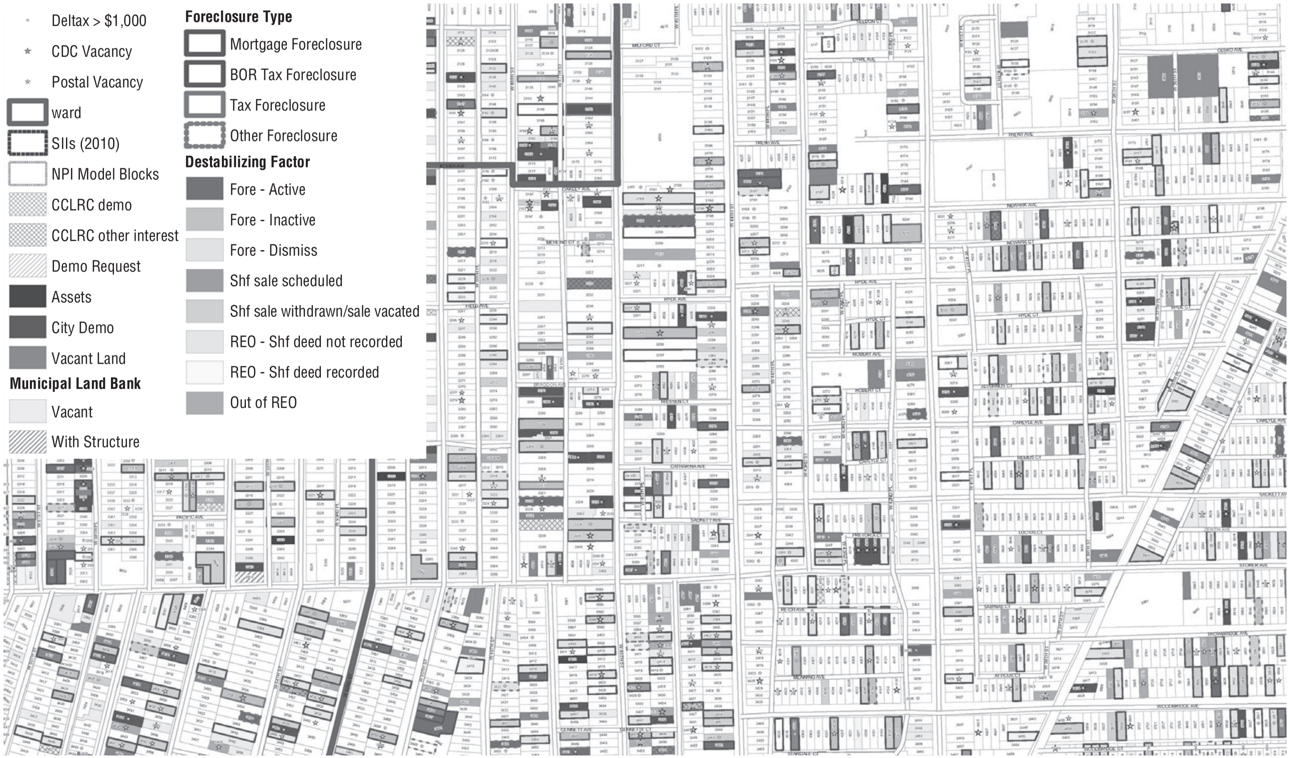
A more demanding level of assembly and analysis is required, however, when communities and city agencies use data to make decisions concerning individual properties within neighborhoods. The need for such a data-driven approach has been recognized for many years, but awareness broadened with the implementation of one of the federal government's first responses to the foreclosure crisis: the Neighborhood Stabilization Program, enacted in 2008 (Joice 2011). This program's funding could be used by localities for a variety of actions affecting properties, including acquisition, rehabilitation, and demolition. The US Department of Housing and Community Development based the first round of grants to states and localities on greatest need, as measured by a variety of housing distress indicators. Recognizing that concentrating housing investment could result in greater impact, the department issued the second round of grants competitively and required grantees to use neighborhood indicators to identify the target neighborhoods.

Limited budgets prevent city- and community-based organizations from dealing with all problem properties, and both neighborhood market conditions and conditions of individual properties should influence their choices. In neighborhoods where market conditions are weak, for example, funding could be wasted by acquiring and rehabilitating too many properties that turn out not to be sustainable in market terms (a higher proportion of demolitions there might have made more sense). Where market conditions are stronger, a less costly strategy (well-focused code enforcement) might be enough to catalyze a trend of self-reinforcing reinvestment. Matching properties and actions in reasonable proportions in a politically difficult and budget-constrained environment represents an extraordinary challenge.

A major problem is that a substantial amount of parcel-specific data is needed to support effective decisionmaking at this level. The Center on Urban Poverty and Community Development at Case Western Reserve University expanded the Cleveland data system discussed at the beginning of this chapter into the most comprehensive information system available for these purposes. Nelson's essay at the end of this chapter presents a case study of this system, named Northeast Ohio Community and Neighborhood Data for Organizing (NEO CANDO), and its many uses.[11]

This system is noteworthy for two reasons. First is the breadth of its content. In addition to more common parcel-level data from assessors and recorders of deeds (e.g., ownership, physical characteristics, and sales prices), the extensive system integrates data from many other sources, including housing code violations, building permits, vacancy status, and foreclosure status. Many of these data are updated weekly. Second is the wide variety of neighborhood and citywide applications supported by NEO CANDO. Since the mid 2000s, the data have been used as the basis for decisions about individual properties within neighborhoods. Various stakeholders, including city officials, representatives of community groups, and NEO CANDO staff, meet to examine parcel-level maps and tables, paying attention to the spatial clustering of conditions as well as the circumstances of individual buildings (see figure 5.1). Discussions of the data in these meetings inform decisions by all participants.

**Figure 5.1.** Neighborhood Stabilization Team Map, Cleveland, Ohio




*Sources:* Various data sources compiled by the Center on Urban Poverty and Community Development, Mandel School of Applied Social Sciences, Case Western Reserve University.

*Note:* A color version of the map is available at http://neighborhoodindicators.org/library/catalog/neighborhood-stabilization-team-map

Public officials and other stakeholders set priorities as to which properties within the neighborhood to acquire, demolish, or rehabilitate. The data also support decisions by community groups working to prevent foreclosures. These groups use listings of properties that the data indicate may be likely foreclosure candidates as the basis for prioritizing door-knocking efforts. For example, they give high priority to a property just entering foreclosure that is next door to one that has just been renovated. The parcel-level information also is used to target actions inside the neighborhood by the city's code enforcement staff and neighborhood groups to put pressure on owners of properties that are not being adequately maintained.

The final aspect that warrants mention is how the use of NEO CANDO spurred collaboration. Over time, participation in the mapping sessions spread to include CDCs, residents, city officials, nonprofit counseling and advocacy groups, realtors, and lenders. Participants have said the fact that they were all using the same data promoted inclusiveness in the process. Reliance on common data meant less disagreement than would otherwise have been the case. Decisions, for example, about which properties to rehabilitate and which to demolish were less controversial because all participants knew the reasoning and facts behind the choices.

## Applications in Large-Scale Comprehensive Initiatives: Examining the Potential for Performance Management

The largest comprehensive community development programs in operation over the past decade were those managed by the Local Initiatives Support Corporation (LISC): its 16 neighborhood New Communities Program (NCP) in Chicago and its Building Sustainable Communities (BSC) initiative, which has operated in just over 100 neighborhoods in 24 other metro areas (Walker et al. 2010).[12] The BSC model calls for coordinated investment across five programmatic domains: housing and real estate, connection to regional economies, family income and wealth generation, education, and public health and safety (Walker et al. 2010).

Key features of both of these initiatives include (1) extensive and continuing community organizing, (2) engaging residents in the preparation of a community quality-of-life plan, (3) enduring community partnerships, and (4) active intermediation across sectors and between the neighborhood and the broader system of support. In each participating neighborhood, a lead agency (an indigenous CDC or other nonprofit) is selected to mobilize other neighborhood organizations and collaborate with them to carry out the plan. This partnership (or "platform"), which is key in attracting resources as well as implementing the work, emphasizes the importance of developing the negotiation skills of all parties (Bridgespan Group 2012, 44).

NCP had no requirements for use of data by lead agencies in the early phases, and it appears that the reliance on data was not uniform or extensive. Neighborhood groups took advantage of available statistics that they considered useful in preparing their quality-of-life plans. As part of the interim evaluation of NCP conducted around the middle of the implementation period, MDRC performed an analysis that influenced the way initiative participants saw the neighborhoods in Chicago. This analysis entailed assembling data on trends in many quality-of-life indicators for both NCP neighborhoods and a comparison group of similar neighborhoods elsewhere in the city (Greenberg et al. 2010, ES1).[13]

A regularly updated and uniform tracking of conditions in most BSC neighborhoods, however, was initiated soon after that initiative began by LISC's national research and assessment office [see overall discussion in Bridgespan Group (2012, 44)]. Data were assembled for the more than 100 neighborhoods from national sources on selected indicators in four areas: housing and real estate, income and wealth, economy and workforce, and demographics.[14] Figure 5.2 is an example of tracking the core indicators of the economy and workforce for eastern North Philadelphia.

Data on all indicators were also assembled for a set of comparison neighborhoods in each city, which were selected because of their similarity to BSC neighborhoods across several topic areas [see methods in Walker, Winston, and Rankin (2009)].[15] In most cases, charts of annual trends go back far enough to show the contrast between trends during the Great Recession and those during the growth period that preceded it. These monitoring reports offer a much richer array of community development–relevant information over a longer timeframe than would have been available anywhere a decade earlier.

Yet more elaborate monitoring in BSC has been implemented with the assistance of local NNIP partners in four cities (Indianapolis, Indiana; Milwaukee, Wisconsin; Minneapolis, Minnesota; and Providence,

**Figure 5.2.** Tracking Core Indicators, Eastern North Philadelphia Neighborhood

## *Core Indicators of Neighborhood Quality*

*Neighborhood* Eastern North Philadelphia

*City* Philadelphia, PA

**Resident Employment**

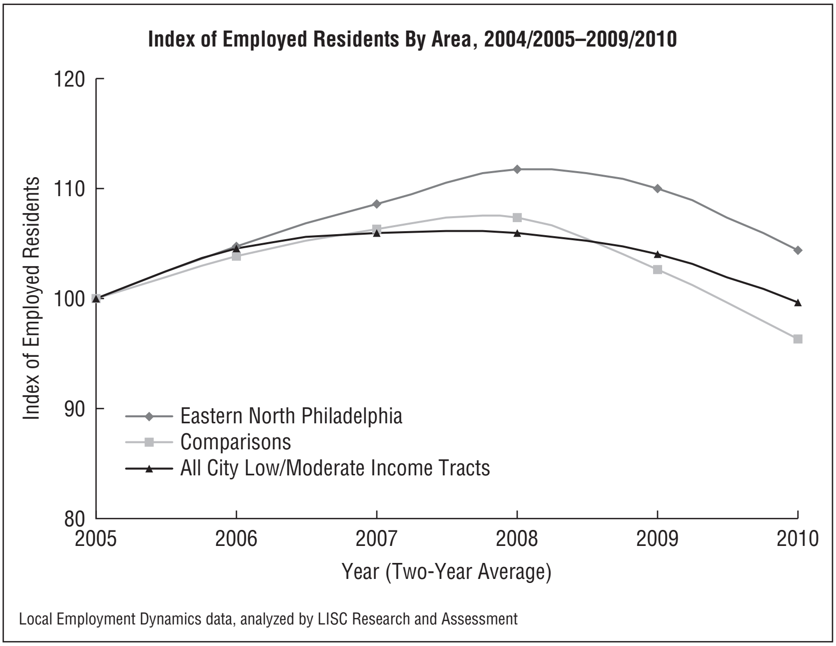
Helping low-income residents get jobs and keep them is one of the most difficult community development challenges. Increased numbers of employed residents are a welcome sign of neighborhood strength. Changes in employment levels, as well as the incomes earned by residents, often are tied to the performance of specific economic sectors, which display different patterns of gain and loss.

***Interpretation of Employed Residents (Chart Below)***

*Overall, the number of employed residents in Eastern North Philadelphia increased slightly between 2005 and 2010, outpacing comparison areas. As noted on page 4, the neighborhood gained population between 2000 and 2010. Based on the pattern among BSC neighborhoods nationwide, Eastern North's gain in employed residents is consistent with the size of its population gains.*

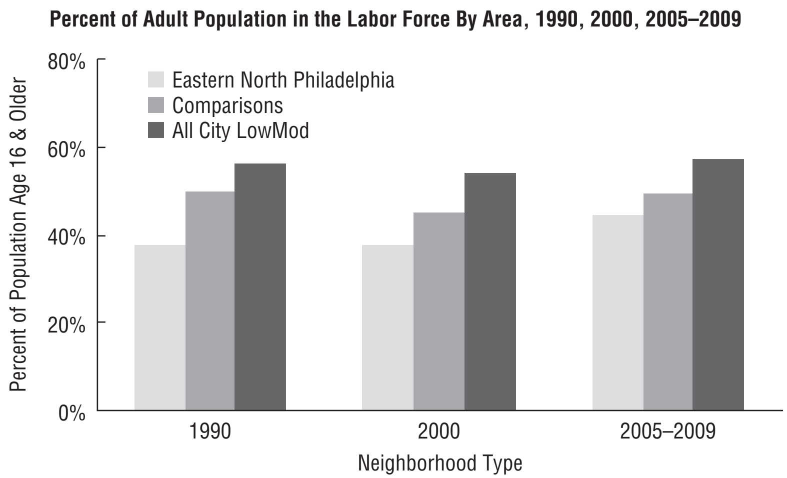
*Target Area Percent of Comparison Area Values:*

| | |
|---|---|
| Most Recent Year | Not Applicable |
| Over Time | + 8% |

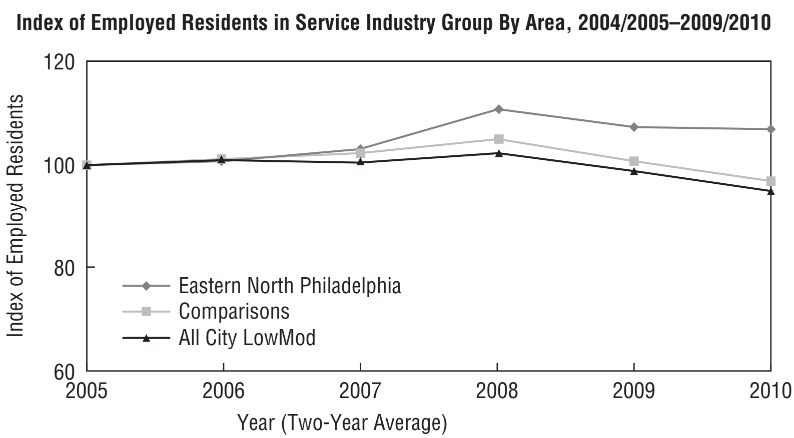



*Source: LISC Neighborhood Monitoring Reports, February 2013.*

## Economy & Workforce




Census 1990, Census 2000 and 2005–2009 American Community Survey 5-Year Estimates, analyzed by LISC Research and Assessment




Local Employment Dynamics data, analyzed by LISC Research and Assessment

Rhode Island covering 15 BSC neighborhoods). In these locations the NNIP partners enhanced the national data with indicators from their local data warehouses (e.g., on crime rates, educational proficiency, health, and home sales prices). This monitoring system has been regularly reviewed by LISC managers in each BSC city from the start, but in 2010 the system also began to be used as the basis for dialogues with lead agencies in the neighborhoods. LISC research staff, local data partners, and lead agency staff review the data and talk about possible implications for the work.

It is important to note the differences between the approach described above and performance measurement systems such as those referred to as "collective impact" reviewed in chapter 2. In the LISC example, the indicators being tracked are a mixture of outcome and context indicators (Land's terms from chapter 4) selected to inform managers about key trends likely to affect their work. But they are not all indicators of goals, and BSC managers do not identify particular targets for each indicator that they would be willing to be held accountable to achieve, as is required in collective impact. In this work, it is critical to recognize the difference between indicators for which changes are driven importantly by outside (e.g., market) forces as well as by the actions of the initiative and indicators on which initiative partners might be expected to have more influence over the results.

Brophy (2011) explicitly recommended the collective impact approach for the community development sector in Milwaukee. Still, implementing collective impact in comprehensive community development would be extremely challenging, much more so than it was in its best-known success story to date, the Strive initiative in education. Kingsley (2011) points out that Strive relied on a comparatively limited number of measures of achievement (e.g., student proficiency scores, high school graduation rates) for which data were fairly readily available, whereas comprehensive community development involves a much broader and more complicated set of goals and, for many of those, data availability remains a serious barrier. For example, some relevant social service programs do not yet maintain machine readable records on the services they provide.

Recognizing the need to move further toward effective performance management, Chicago LISC implemented a new approach to NCP beginning in 2011. Called Testing the Model (TTM), this effort entails applying collective impact–like methods, but only in one domain each (e.g., education or public safety) in 7 of the 16 original NCP neighborhoods. This practice is consistent with our point above that a comprehensive initiative can start with one domain and then, based on what is learned, develop more thorough performance management in other domains over time.

Susana Vasquez, director of LISC Chicago, sees an aim of this effort as helping "neighborhood leaders get excited about using data to improve their work" (Kelleher 2012). TTM begins by developing a logic model of the change process in the selected domain (a theory of change as introduced earlier). In LISC's conception, this model requires explicit statements of expectations about the links between resources, activities, outputs, and outcomes. As one example, the logic model for the TTM effort in the Pilsen neighborhood, titled "The Pilsen Bridge: Pathways for Better Transitions to Kindergarten and High School," describes proposed interventions, paths of influence, and targeted outcomes (figure 5.3).

Given the diversity of aims, LISC recognized that each neighborhood would have to develop its own data.[16] Technical assistance and other supports have been provided to help them assemble the data to monitor their selected outcomes. Many of the performance measures selected could be developed from the records of partner organizations in the neighborhood. But the task of obtaining useful data from the administrative records of city-level agencies was more challenging. For example, outcome measures may require aggregating individual-level records stored in confidential school or service program information systems. Special programming is normally needed to select out records for neighborhood residents and calculate new performance measures for them, as are efforts to assure privacy protections.

TTM is still in process and it is too early to assess results. However, to our knowledge, it is the first effort to try to actually measure such a broad range of results indicators in a comprehensive community development setting. Importantly, this includes measures of status changes for children and families at the neighborhood level from administrative records that have rarely been made operational in any setting. This experience offers a promising model for performance monitoring in similar programs elsewhere. Other signs leading to the same conclusion are noted in our discussions of the Promise Neighborhoods Initiative and integrated data systems later in this chapter.

**Figure 5.3.** Theory of Development from the Pilsen Bridge Neighborhood in Chicago, Illinois

## THE PILSEN BRIDGE:

### Pathways for Better Transitions to Kindergarten and High School

**If . . .** we improve student programming and parent participation at two critical transition points in the educational continuum—from early childhood education to kindergarten, and from middle school to high school—and help students and their families make successful transitions from each of these levels to the next,

**Then . . .** we will see increases in student academic and social-emotional preparedness, matriculation rates and school performance,

**Which will lead to . . .** improved student achievement in elementary schools, increases in high school graduation rates, and measurable gains towards preparing our students to graduate from college and participate in the work world,

**And over the longer term . . .** we will create a stable, safer and economically thriving community.

| INTERVENTIONS | OUTCOMES | KEY PARTNERS |
|---|---|---|
| 1. Improve the quality and quantity of early-childhood development programs.<br>2. Create parent-child literacy programs in local pre-K and elementary schools.<br>3. Connect off-track students to community-based programs that will improve their transitions to local high schools.<br>4. Align curricula and create teacher exchanges among receiving high school and feeder elementary and middle schools. | 1. More students are prepared when entering kindergar-ten.<br>2. More students are on track for reading by 3rd grade.<br>3. Parents have an increased comfort level in schools.<br>4. More students are on track for 9th grade.<br>5. High school graduation rate improves.<br>6. High school dropout rate falls.<br>7. Curricula are aligned from middle to high school.<br>8. Teachers are better prepared to assist in successful transitions to high school. | • **The Resurrection Project (lead)**<br>• **Chapin Hall, U of C (data intermediary)**<br>• Early childhood–El Hogar del Nino, El Valor, Chicago Commons<br>• CPS Elementary Schools–Pilsen Academy, Jungman, Cooper, Finkl<br>• CPS Middle Schools–Ruiz, Perez, Orozco<br>• CPS High School–Juarez<br>• Instituto del Progreso Latino, Casa Juan Diego, Gads Hill Center, National Museum of Mexican Art |

*Source:* LISC Chicago.

## New Insights on Neighborhood Change from Making Connections

Making Connections was a major community-building initiative sponsored by the Annie E. Casey Foundation in the 2000s. Its central goal was to improve outcomes for vulnerable children living in tough neighborhoods. Its strategy was based on the belief that the best way to achieve that goal was to strengthen their families' connections to economic opportunity, positive social networks, and effective services and supports.

As of 2002, the initiative covered ten cities: Denver, Colorado; Des Moines, Iowa; Hartford, Connecticut; Indianapolis, Indiana; Louisville, Kentucky; Milwaukee, Wisconsin; Oakland, California; Providence, Rhode Island; San Antonio, Texas; and White Center (a community south of Seattle, Washington). One or more low-income neighborhoods were selected as the focus of activity in each site (average population of 40,000 per site). Program activities were directed by Casey site managers for most of the period, but responsibility was shifted to local management entities in the later years.[17]

The most noteworthy contribution of Making Connections related to advancing the use of neighborhood data was its program of cross-site surveys.[18] Three waves of surveys were conducted, with in-person interviews of sizable samples of resident households in the Casey neighborhoods. Depending on the city, the first wave was conducted between 2002 and 2004, the second between 2005 and 2007, and the third between 2008 and 2009.[19] In all waves, the surveys covered a wide range of topics including demographics, child well-being, household finances, neighborhood attachment, and perceptions of local services and neighborhood conditions.

The wave 1 survey entailed interviews with samples of around 800 households per city. In wave 2, the interviews were conducted with the residents of the same sample of housing units involved in the earlier wave but, in addition, with families with children who were interviewed in wave 1 that had moved to another housing unit. The method was repeated in wave 3. This feature is unique among neighborhood-focused surveys. For the first time, analysts were able to describe not only the net change that occurred in the neighborhood (e.g., the change in the employment rate), but also the degree to which the trend was caused by changes in the circumstances of residents who stayed in the neighborhood versus the differential characteristics of in-movers and out-movers.

The Making Connections surveys have been the basis for many research products.[20] Many of the completed products show how conditions changed in these neighborhoods over the survey period. One important report, for example, presented data on how the wealth of low-income neighborhood families shifted before and after the onset of the Great Recession (Hendey, McKernan, and Woo 2012).[21]

Two strands of the Making Connections findings on the process of neighborhood change fundamentally altered our understanding of the way neighborhood indicators and performance measures need to be examined. The first and most important concerns residential mobility. Coulton, Theodos, and Turner (2009) found a high rate of residential mobility in Making Connections neighborhoods; 61 percent of the families with children interviewed in the wave 1 survey had moved by wave 2. This translates into an annual mobility rate of 28 percent (Kingsley, Jordan, and Traynor 2012).

The idea of high mobility disturbs community developers working to build social capital in neighborhoods. Although research showing higher mobility rates for low-income groups compared with households with higher incomes has been available for some time,[22] the fact has not been much discussed in the community development literature. The Making Connections research, however, performed for an initiative whose central purpose was community improvement, seems to be gaining wider recognition in that field. Kubisch et al. (2010, 140), for instance, states "[t]he next generation of community change efforts must take up the challenge of developing good theories of change that reflect this new understanding of . . . mobility dynamics."

Actually, the real news from the Making Connections research was not about the overall extent of mobility, but about its composition. Coulton, Theodos, and Turner (2009) performed a cluster analysis that divided the family household movers into three groups:

- *Up and out moves* (30 percent of all movers). Households with fairly high incomes (median $28,000) who sought a better home or neighborhood. They moved the longest distances (median 5.8 miles) and were generally satisfied and optimistic about their new neighborhoods.
- *Nearby attached movers* (25 percent of movers). Households with much lower incomes (median $15,000) who typically moved a short distance (median 1.1 miles). Their moves were more often

due to life-cycle factors (e.g., a new baby meant they needed an apartment with more room) than dissatisfaction with their old house or neighborhood. They generally had positive views of their neighborhood and new unit post-move.

- *Churning movers* (45 percent of movers). Also households with very low incomes (median $14,000) who moved short distances (median 1.7 miles). They generally viewed their neighborhoods as unsafe and not good places to raise children. It seems likely that many of them felt forced to move because of financial stress or problems with their rental housing arrangements.

These findings significantly alter the way community developers should regard mobility. First, there is no need to be disheartened by the idea that a large share of the people who community initiatives are trying to help will move away, as (and this is the major surprise) the majority of the movers relocate within or near their original neighborhood. The nearby attached and churning movers account for 70 percent of the Making Connections moves. Second, recognizing that some moves are positive for families, community developers can focus energy on reducing the potentially problematic ones, the churning moves that represent residential instability that can be very costly for children in particular (Kingsley, Jordan, and Traynor 2012).

Another aspect of mobility needs to be taken into account when interpreting data on trends in community well-being. Making Connections researchers recognized that neighborhood economic indicators change due to differentials in the incomes of in-movers and out-movers as well as to changes in the incomes of the residents who do not move. For example, a neighborhood's poverty rate will go down if a large number of poor residents move out, the average poverty rate of the in-movers is the same as the initial neighborhood average, and the incomes of those who stay do not change at all. The researchers calculated the implied components of change in the poverty rates of Making Connections neighborhoods between the wave 1 and wave 2 surveys. They found that explanations varied across neighborhoods but, overall, "changes in poverty occurred primarily through mobility, not because of changing circumstances for stayers" (Coulton, Theodos, and Turner 2012, 75).

In addition to the results on mobility, Making Connections research yielded new findings on another topic that also alters how community improvement initiatives understand outcomes. Theodos, Coulton, and

Budde (2014) examined the pattern of school attendance in the 10 cities.[23] Their findings confirm a shift that has been known to be under way for some time as more and more school districts adopt school choice policies. In every one of the Making Connections neighborhoods, large shares of the elementary school students who live in the neighborhood attend schools outside the neighborhood (and, presumably, the schools inside the neighborhood serve many students who live elsewhere). This means, for instance, that improvements in the academic proficiency scores of neighborhood children may be explained as much by advances made by actors outside the neighborhood as by the programmatic efforts within the neighborhood. This research on schools reinforces awareness of a reality that exists for many other services in community initiatives (e.g., financial counseling, job placement services)—namely, that it is challenging to precisely align service populations and neighborhood boundaries.

In most communities, in-depth, longitudinal data on mobility and service area differences are not available.[24] The lack of complete data, however, does not mean that these issues cannot be thoughtfully considered in assessment and decisionmaking in community development. When stakeholders review trends in neighborhood outcomes, as pointed out above, the influence of mobility is too important to be ignored, and it does not have to be. National and administrative data sources, such as the American Community Survey, can offer clues about residential movement, and focus groups and interviews can reveal some understanding of its composition. Even with such imperfect knowledge, fruitful discussions are possible exploring the implications of what initiative leaders know from qualitative and quantitative sources in relation to an initiative's logic model.

Furthermore, there are prospects that richer data on neighborhood change dynamics may become available at reasonable cost in the future, not from surveys, but from more effective exploitation of administrative datasets. There were hints of this in the TTM experience, and there are more in the early implementation of the Promise Neighborhoods program, which we review next.

## Intensive Data Use in Program Planning and Implementation: Promise Neighborhoods

As noted in chapter 2, Promise Neighborhoods is a US Department of Education program modeled after the well-known Harlem Children's Zone [documented by Tough (2008)]. The founder of that effort, Geoffrey Canada, well understood the importance of place. He focused on building a "continuum of solutions" that would guide the development of children from birth, through schooling, and into successful careers, and purposefully did so in one neighborhood such that the synergy of in-community relationships could enhance results.

Promise, like others of the Neighborhood Revitalization Initiative programs, seeks integrative solutions through work that breaks down programmatic silos and deals with neighborhoods and families in a holistic manner. Furthermore, these programs urge or require planning based on analysis of trends in neighborhood conditions and performance management to monitor progress and make corrections as needed. One of the five key elements for an effective neighborhood revitalization strategy is maintaining a results focus supported by data:

> A results framework presents a strategy for achieving specific objectives, helping to focus multiple stakeholders on a common goal, and creates a dataset for the community policymakers and other stakeholders to use to measure progress over time. Data should not only measure population-level outcomes but should also drive the development of the other elements identified in this report . . . data [are] a critical tool for building cross-agency accountability systems and tracking progress against desired results. (White House 2011, 7)[25]

The emphasis on data in Promise Neighborhoods was well publicized when the federal program was in the planning stages and groups in many cities appeared ready to respond. One example is the District of Columbia Promise Neighborhoods Initiative (DCPNI), which began work in 2008. Irasema Salcido, founder of the Cesar Chavez Public Charter Schools, saw that the academic skills of many students entering her middle and high schools in Kenilworth-Parkside were well below grade level. She convened a steering committee to adapt the comprehensive approach of the Harlem Children's Zone to address the challenges faced by the children and the distressed community in which they lived. A collaborative was formed to develop and carry out the program that involved the Cesar Chavez Public Charter Schools, traditional District public elementary schools, government agencies, and a number of local service providers, accompanied by active resident engagement (Comey et al. 2012).

By the time it received one of the first-round federal planning grants in 2010, DCPNI had already engaged a data partner: the Urban Institute, also home to the local NNIP partner, NeighborhoodInfo DC. The data partner was to be an unusually active participant in the planning process.

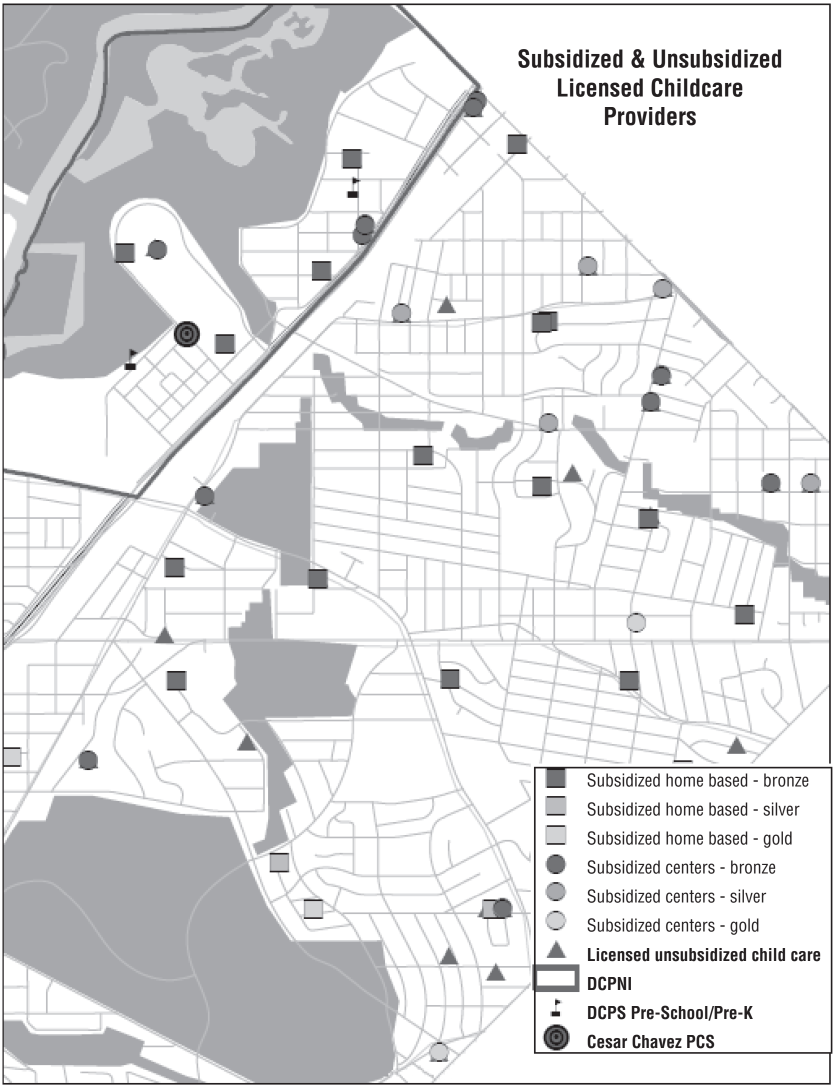
This process began with their development of a needs assessment and segmentation analysis[26] in the period from October 2010 to July 2011 (Popkin et al. 2011). This work included (1) a comprehensive analysis of neighborhood conditions and trends using national and local datasets, (2) focus groups with adult and teenage residents, (3) resident retreats, (4) teacher interviews, and (5) a school climate survey of middle school and high school students.

Most noteworthy, however, was the role played by data partner staff in program planning. The planning process was conducted by eight results-driven working groups that corresponded to the goals that had been selected for DCPNI. Data partner staff were embedded as participants in each of the groups. They would select relevant data to present at meetings and then facilitate the discussion with the other participants about the implications for program design. These "dialogues with data" were credited as having an important influence on the plans that emerged. As an example, one of the goals was ascertaining the "percent and number of young children in center-based or formal home-based early learning programs." The Urban Institute mapped the center-based child care locations, their capacities, and their quality ratings to identify the current supply (figure 5.4). Although about half of all children age 4 and younger were enrolled in formal early child care, the early child care providers were largely rated as low quality. This finding led to the opening of a new high-quality child care center and efforts to improve the capacities of other providers.[27]

In the next phase of Promise nationally—implementation and performance management—the use of data is yet more intensive. A publication has been developed that offers guidance on data systems and indicators for Promise implementation grantees (Comey et al. 2013).[28] A critical federal requirement is that all grantees monitor trends for 15 indicators related to the results the initiative is intended to achieve. These are also the measures that the Department of Education has identified to hold the program accountable under the Government Performance and Results Act (GPRA). Examples are as follows:[29]

- *GPRA 2.* Number and percentage of three-year-olds and children in kindergarten who demonstrate at the beginning of the program or school year age-appropriate functioning across multiple domains of early learning

**Figure 5.4.** Licensed Childcare Provider Map of the Kenilworth-Parkside Neighborhood, District of Columbia




*Source:* Urban Institute mapping of data from the Office of the State Superintendent of Education, Division of Early Childhood Education, 11/16/10.
*Note:* A color version of the map is available at http://www.neighborhoodindicators.org/library/catalog/nnip-and-place-based-initiatives

- *GPRA 4.* Number and percentage of students at or above grade level according to state mathematics and English language arts assessments in at least grades 3rd through 8th and once in high school
- *GPRA 10.* Number and percentage of students who feel safe at school and traveling to and from school

In order to report the GPRA measures, Promise grantees must regularly collect neighborhood-level data from national and local data files and special surveys on school climate and various neighborhood conditions. However, a more formidable challenge is for them to maintain data on individual children in a case management system. The data on services are collected for each child in the program and typically cover demographics, the type and dates of services provided, level of service (e.g., number of hours in counseling), and indicators of achievement (e.g., grade in a course that has been completed).[30]

Although case management data are required for only two of the GPRA measures, Promise grantees are nonetheless encouraged to develop a case management system because of the importance of such systems to effective program management and outcomes. It is very difficult, for example, for case managers to offer sensible advice about next steps in supporting individual children and families in the program without considerable knowledge about their circumstances and their prior service histories.

Much of the challenge for systems development, however, arises from the requirement that Promise Neighborhoods coordinate and track a number of services delivered by separate local providers.[31] This implies the need to integrate records that are maintained by separate providers on individual children and their families. It might involve, for example, integrating data about a student from school records (e.g., proficiency scores, absenteeism, school mobility) with data on the same child from child welfare agencies, health care providers, and afterschool programs.

Credible summary information on what this type of program achieves requires summarizing data on what happens to individuals, but the task of data integration is obviously difficult, especially if attempted on a small scale. This task has certainly been a formidable challenge for the LISC's TTM effort in Chicago. Perhaps more important than the technical difficulties are the steps that must be taken to maintain high standards regarding confidentiality. This work inevitably presses the boundaries set under the HIPAA and FERPA legislation noted in chapter 2. Even so, as documented in chapter 3, the number of efforts to develop jurisdiction-wide integrated data systems (IDS) along these lines is expanding rapidly. There appears to be growing recognition that important net benefits are possible from cross-agency data sharing even while paying serious attention to the risks.

The Promise Neighborhoods implementation grantees are developing better data capabilities to meet their management and reporting needs within their selected neighborhoods. However, the sponsors recognize the importance of monitoring mobility:

> The most challenging group to track will be children who lived in the Promise Neighborhood after the start of the program but have since moved away . . . this mobility itself may be an outcome. If these children are not a part of the *data universe* the Promise Neighborhood will not be able to track this kind of success. (Comey et al. 2013, 40)

Telling the stories of what happened to children who lived or were served in the neighborhood, whether they were present for a part or for all of the initiative's duration, is critical to an accurate interpretation of the results of the program, but it will present an ongoing challenge for the initiative leaders.

## Conclusions

We started this chapter talking about information systems with a modest list of indicators for neighborhoods as a whole that were updated most often annually. This capacity was regarded as an important data breakthrough in the early 1990s. We end the chapter noting that at least some institutions are now able to monitor neighborhood change on almost a real-time basis with much more detailed information on the changing circumstances of individual properties. Our first conclusion from this chapter, then, is that tremendous progress has been made with respect to timeliness and the depth of the data available.

Second, there have been impressive instances in which the new data have been applied to great advantage in community decisionmaking. These have included applications that have more sharply diagnosed community problems and opportunities, designed effective actions in response, and managed performance to assure successful implementation. And they have also included cases in which the data motivated support for sound policies and provided a deeper understanding of the processes of neighborhood change.

These advanced applications, however, are the exception rather than the rule. Even most of these examples are still experimental and are not yet built into the routine processes of local decisionmaking. Our third conclusion is that there is still much work to be done and important barriers to be overcome before the full promise of community information can be attained, before the key institutions of local governance (public and private) in most of America can truly be considered to be data driven. In chapter 8, we offer ideas about steps that might be taken to more fully take advantage of the potential of the community information field.

## NOTES

1. A more complete version of this story, and of the history of community development overall, is provided in Von Hoffman (2012).

2. These range from the Community Development Block Grant program, enacted in 1974, to the establishment of the Community Development Financial Institutions (CDFI) Fund at the Treasury Department in the early 1990s. Also ultimately important was the regulatory force of the Community Reinvestment Act (CRA) which was designed to overcome bank resistance to lending in inner-city neighborhoods. See von Hoffman (2012) and Rohe (2009).

3. One influential example is the Comprehensive Community Revitalization Program in the South Bronx launched in 1992 (Miller and Burns 2006). A number of CDC practitioners had favored a more comprehensive approach before the 1990s, but they were in the minority.

4. No formal studies were conducted of how the data were used in CCBI. However, using interviews of staff participants, Milligan et al. (1998) document and analyze the process by which CCBI participants developed their "theories of change" to guide the work, uncovering both benefits and tensions.

5. Terri J. Bailey amplified this theme in "Advances in the Science and Practice of Community Indicators," a presentation for the Community Quality of Life Conference, Reno, Nevada, March 11, 2004.

6. See www.abcdinstitute.org for publications from the Asset-Based Community Development Institute.

7. Discussed further in the section on Promise Neighborhoods.

8. Samantha Teixeira and John Wallace, "Data-Driven Organizing: Partnerships for Community Change in the Homewood Children's Village," a presentation at Pittsburgh Neighborhood and Community Information System Annual Users' Conference, June 11, 2010.

9. For explanations of the broad foundations and effects of the foreclosure crisis, see Gramlich (2007), Herbert and Apgar (2009), and Carr (2008).

10. Research on neighborhood effects of foreclosures is summarized in Immergluck and Smith (2006) and Kingsley, Smith, and Price (2009).

11. Aspects of the NEO CANDO system and some of its applications have also been documented in Sand and Bush (2011) and Pettit and Kingsley (2011). Data from the system have also been used to support a number of citywide applications, as discussed further in Nelson's essay at the end of this chapter.

12. NCP's full-scale operation ran for 10 years (2003–2012). The program received $47 million in grants from the MacArthur Foundation and, LISC reports, an equal amount from other major foundations, which in turn leveraged $540 million more in neighborhood investments (LISC Chicago 2011). BSC was launched in 2007, and according to the Bridgespan Group (2012), LISC was investing at an annual rate of $107 million across all BSC neighborhoods in 2011.

13. MDRC is continuing its evaluation of NCP. A recent report in its series covers broader "policy change and systems change" activity in the program (Chaskin and Karlström 2012).

14. The sources include the US Decennial Census, the American Community Survey, the Home Mortgage Disclosure Act, Local Employment Dynamics, IRS tax statistics, US Postal Service vacant addresses, and National Center for Education Statistics datasets. These sources are all available to the public and are described in chapter 3. LISC also relies on proprietary data on mortgage status from LPS Analytics for these reports.

15. Clearly, the comparison neighborhood approach does not meet the same standards for evaluation as would a random control trial (see discussion of evaluation in chapter 7), but as that alternative is not available here it would seem that determining where the performance of a BSC neighborhood is notably above or below that of its comparison neighborhoods should be a valuable indicator for initiative managers.

16. In an interview, Vasquez indicated that LISC plans to revisit what administrative data (such as school mobility) or other local data (such as connecting individuals to community institutions) could be tracked across plans to anchor a shared performance management system.

17. The foundation has written an overall review of the Making Connections experience that explains program operations and other features in greater detail (Annie E. Casey Foundation 2013).

18. The survey design team was composed of staff from NORC at the University of Chicago and the Urban Institute. NORC was responsible for survey administration. Feister (2011, 13–16) reviews issues faced in survey design and how they were addressed. The final in-person household survey instrument contained 180 questions and took 45 minutes to complete. The overall response rate was 75 percent.

19. Three of the 10 cities (Hartford, Milwaukee, and Oakland) did not have wave 3 surveys.

20. The central survey team included researchers from the Urban Institute, Case Western Reserve University, and NORC, as well as the Annie E. Casey Foundation. The website (http://mcstudy.norc.org) lists and provides access to these products, explains and documents the survey in greater detail, and explains how researchers can access the data.

21. The authors found, for example, that both asset and debt amounts increased between 2005–06 and 2008–09, but asset and debt levels remained lower for vulnerable

families, and low-income families disproportionately lost equity during the crisis. Yet even in 2008–09, home equity remained a substantial component of wealth for nearly half the families who were homeowners (more than four times as much as families had in savings).

22. For example, see discussion in Theodos (2012).

23. The main aim of this study was to shed light on how the circumstances surrounding school changes influence whether children attend better- or worse-performing schools over time. The results were disappointing. The authors found that switching to poorly performing schools is most frequent among families moving short distances in response to financial distress or household compositional changes. Only when relocation takes families outside the originating school district do they see reliable gains in terms of their children attending schools with higher-performing peers.

24. Surveys like those the Annie E. Casey Foundation supported in Making Connections are very expensive. Feister (2010) reports that the total cost of implementing the cross-site surveys was $20.2 million. Thus the average cost of implementing one wave in the selected neighborhoods in one city (average 40,000 population) was about $750,000.

25. The five elements are (1) resident engagement and community leadership, (2) developing strategic and accountable partnerships, (3) maintaining a results focus supported by data, (4) investing in and building organizational capacity, and (5) aligning resources to a unified and targeted impact strategy.

26. The segmentation analysis identifies, and presents more detailed information on, subpopulations judged to be most in need.

27. G. Thomas Kingsley, Kathryn L.S. Pettit, and Jennifer Comey, "NNIP and Place-Based Initiatives," a presentation to Grantmakers for Effective Organizations, Washington, DC, July 19, 2012.

28. From fiscal years 2010 through 2012, the federal government awarded 46 Promise Neighborhood planning grants and 12 implementation grants. For a full listing see http://www2.ed.gov/programs/promiseneighborhoods/index.html.

29. All are defined and explained in Comey et al. (2013, Chapter 4).

30. See discussion in Comey et al. (2013, Chapter 5).

31. This is unlike the Harlem Children's Zone, where the multiple services provided are all overseen by the initiative's central staff.

## REFERENCES

Ahlbrandt, Roger S., Jr., and Paul C. Brophy. 1975. *Neighborhood Revitalization.* Lexington, MA: Lexington Books.

Annie E. Casey Foundation. 2013. *Lessons Learned: Community Change Lessons from Making Connections.* Baltimore: Annie E. Casey Foundation.

Biglan, Anthony, Christine Cody, William Aldridge II, Alexis Dubroski, and Jean Kjellstrand. 2011. "The Promise Neighborhoods Initiative: Improving Developmental Outcomes through Comprehensive Interventions." *Journal of the Institute for Comprehensive Community Development* 2 (2): 25–33.

Bridgespan Group. 2012. *The Next Generation of Community Revitalization: A Work in Progress.* Boston: Bridgespan Group, Inc.

Brophy, Paul C. 2011. *Community Development in Milwaukee: An Assessment and Recommendations.* Baltimore: Brophy and Reilly, LLC.

Buchanan, T. K., Phyllis G. Betts, Jackson Gilman, and Robert Brimhall. 2010. *Neighborhood by Neighbor: A Citywide Problem Property Audit.* Memphis, TN: Center for Community Building and Neighborhood Action, University of Memphis.

Carr, James H. 2008. "Responding to the Foreclosure Crisis." *Housing Policy Debate,* 18(4): 837–860.

Chaskin, Robert J., Prudence Brown, Sudhir Venkatesh, and Avis Vidal. 2001. *Building Community Capacity.* New York: Aldine de Gruyter.

Chaskin, Robert J., and Mikael Karlström. 2012. *Beyond the Neighborhood: Policy Engagement and Systems Change in the New Communities Program: Executive Summary.* New York: MDRC. June.

Cleveland Foundation Commission on Poverty. 1992. *The Cleveland Community Building Initiative: The Report and Recommendation of the Cleveland Foundation Commission on Poverty.* Cleveland, OH: Mandel School of Applied Social Sciences, Case Western Reserve University.

Comey, Jennifer, Peter A. Tatian, Lesley Freiman, Mary K. Winkler, Christopher Hayes, Kaitlin Franks, and Reed Jordan. 2013. *Measuring Performance: A Guidance Document for Promise Neighborhoods on Collecting Data and Reporting Results.* Washington, DC: Urban Institute.

Comey, Jennifer, Molly M. Scott, Susan J. Popkin, and Elsa Falkenburger. 2012. *Bringing Promise to Washington DC: The DC Promise Neighborhood Initiative,* Washington DC: Urban Institute.

Coulton, Claudia J. 1995. "Using Community Level Indicators of Children's Well-Being in Comprehensive Community Initiatives." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Context,* edited by James P. Connel, Anne C. Kubisch, Lisbeth Schorr, and Carol H. Weiss. Washington, DC: Aspen Institute.

Coulton, Claudia J., Brett Theodos, and Margery A. Turner. 2009. *Family Mobility and Neighborhood Change: New Evidence and Implications for Community Initiatives.* Washington, DC: Urban Institute.

Coulton, Claudia J., Brett Theodos, and Margery A. Turner. 2012. "Residential Mobility and Neighborhood Change: Real Neighborhoods Under the Microscope." *CityScape: A Journal of Policy Development and Research* 14 (3): 55–90.

Cowan, Jake. 2007. *Stories: Using Information in Community Building and Local Policy.* Washington, DC: Urban Institute.

Dorfman, Diane. 1998. *Mapping Community Assets Workbook.* Portland, OR: Northwest Regional Education Laboratory.

Duke, Elizabeth. 2012. "Foreword: Building Sustainable Communities." In *Investing in What Works for America's Communities: Essays on People, Places and Purpose,* edited by Nancy O. Andrews and David J. Erickson. San Francisco: Federal Reserve Bank of San Francisco and Low-Income Investment Fund.

Feister, Leila. 2011. *Measuring Change While Changing Measures: Learning In, and From, the Evaluation of Making Connections.* Baltimore MD: The Annie E. Casey Foundation.

Gittell, Ross, and Avis Vidal. 1998. *Community Organizing: Building Social Capital As a Development Strategy.* Thousand Oaks, CA: Sage Publications.

Gramlich, Edward M. 2007. *Subprime Mortgages: America's Latest Boom and Bust.* Washington, DC: The Urban Institute Press.

Green, Gary P., and Anna L. Haines. 2002. *Asset Building and Community Development.* Thousand Oaks, CA: Sage.

Greenberg, David, Nandita Verma, Ken-Nicole Dillman, and Robert Chaskin. 2010. *Creating a Platform for Sustained Neighborhood Improvement: Interim Findings from Chicago's New Communities Program.* New York: MDRC. February.

Hendey, Leah, Signe-Mary McKernan, and Beadsie Woo. 2012. *Weathering the Recession: The Financial Crisis and Family Wealth Changes in Low-Income Neighborhoods.* Baltimore: Annie E. Casey Foundation.

Herbert, Christopher E., and William C. Apgar. 2009. "Interim Report to Congress on the Root Causes of the Foreclosure Crisis." Washington: U.S. Department of Housing and Urban Development (March).

Immergluck, Dan, and Geoff Smith. 2006. "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values." *Housing Policy Debate* 17(1): 57–79.

Joice, Paul. 2011. "Neighborhood Stabilization Program." *CityScape: A Journal of Policy Development and Research* 13 (1): 135–41.

Kelleher, Maureen. 2012. "Thinking through Evaluation." *Institute of Comprehensive Community Development Newsletter.* April 16.

Kingsley, G. Thomas, ed. 1999. *Building and Operating Neighborhood Indicator Systems.* Washington, DC: Urban Institute.

Kingsley, G. Thomas. 2011. "Data and Comprehensive Community Development: Unlocking the Potential." *Journal of the Institute for Comprehensive Community Development* 2 (1) July: 10–17.

Kingsley, G. Thomas, Claudia J. Coulton, Michael Barndt, David S. Sawicki, and Peter Tatian. 1997. *Mapping Your Community: Using Geographic Information to Strengthen Community Initiatives.* Washington, DC: US Department of Housing and Urban Development. October.

Kingsley, G. Thomas, Audrey Jordan, and William Traynor. 2012. "Addressing Residential Instability: Options for Cities and Community Initiatives." *CityScape: A Journal of Policy Development and Research* 14 (3): 161–84.

Kingsley, G. Thomas, Robin Smith, and David Price. 2009. *The Impacts of Foreclosures on Families and Communities.* Washington: The Urban Institute.

Kingsley, G. Thomas, Joseph McNeely, and James O. Gibson. 1997. *Community Building Coming of Age.* Washington, DC: Development Training Institute and Urban Institute.

Kretzmann, John P., and John L. McKnight. 1993. *Building Communities from the Inside Out: A Path toward Finding and Mobilizing Community Assets.* Evanston, IL: Center for Urban Affairs and Policy Research, Northwestern University.

Kubisch, Anne C., Patricia Auspos, Prudence Brown, Robert Chaskin, Karen Fulbright-Anderson, and Ralph Hamilton. 2002. *Voices from the Field II: Reflections on Comprehensive Community Change.* Washington, DC: Aspen Institute.

Kubisch, Anne C., Patricia Auspos, Prudence Brown, and Tom Dewar. 2010. *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts.* Washington, DC: Aspen Institute.

Kubisch, Anne C., Prudence Brown, Robert Chaskin, J. Hirota, Mark Joseph, Harold Richman, and M. Roberts. 1997. *Voices from the Field: Learning from the Early Work of Comprehensive Community Initiatives.* Washington, DC: Aspen Institute.

LISC Chicago. 2011. *Report to Donors, 2011.* Chicago: LISC Chicago.

LISC. 2013. Neighborhood Monitoring Reports, February. Unpublished report.

Miller, Anita, and Tom Burns. 2006. *Going Comprehensive: Anatomy of an Initiative That Worked: CCRP in the South Bronx.* Philadelphia, PA: OMG Center for Collaborative Learning. December.

Milligan, Sharon, Claudia Coulton, Peter York, and Ronald Register. 1998. "Implementing a Theory of Change Evaluation in the Cleveland Community Building Initiative." In *New Approaches to Evaluating Community Initiatives: Volume 2: Theory, Measurement, and Analysis,* edited by Karen Fulbright-Anderson, Anne C. Kubisch, and James P. Connel (45–85). Washington, DC: Aspen Institute.

National Congress for Community Economic Development. 1989. *Against All Odds: The Achievements of Community-Based Development Organizations.* Washington, DC: National Congress for Community Economic Development.

National Congress for Community Economic Development. 2005. *Reaching New Heights: Trends and Achievements of Community-Based Development Organizations.* Washington, DC: National Congress for Community Economic Development.

Orzag, Peter R., Melody Barnes, Adolfo Carrion, and Lawrence Summers. 2009. "Developing Effective Place-Based Policies for the FY 2011 Budget." Memorandum. Washington, DC: Office of Management and Budget, Domestic Policy Council, Office of Urban Affairs, and National Economic Council. August 11.

Paloma, Marjorie. 2012. "The Intersection of Health, Philanthropy and Housing." *Shelterforce* Spring. Accessed February 7, 2014. http://www.shelterforce.org/article/2768/the_intersection_of_health_philanthropy_and_housing/.

Pastor, Manuel, and Margery Austin Turner. 2010. *Reducing Poverty and Economic Distress after ARRA: Potential Roles for Place-Conscious Strategies.* Washington, DC: Urban Institute. April.

Pettit, Kathryn L. S., and G. Thomas Kingsley. 2011. "Framework: The New Potential for Data in Managing Neighborhood Change." In *Putting Data to Work: Data-Driven Approaches to Strengthening Neighborhoods.* Washington, DC: Federal Reserve Board of Governors. December.

Popkin, Susan, Jennifer Comey, Molly Scott, Elsa Falkenberger, Chantal Haley, and Amanda Mireles. 2011. *DC Promise Neighborhood Initiative: Needs Assessment and Segmentation Analysis.* Washington, DC: Urban Institute.

Promise Neighborhoods Institute. 2011. *What Is a Promise Neighborhood?* Oakland, CA: Promise Neighborhoods Institute. http://www.promiseneighborhoodsinstitute.org/What-is-a-Promise-Neighborhood.

Rohe, William M. 2009. "From Local to Global: One Hundred Years of Neighborhood Planning." *Journal of the American Planning Association* 75 (2) Spring: 209–30.

Sand, Peggy, and Phillip Bush. 2011. *Market Data-Driven Stabilization: A Case Study of Cleveland's NEO CANDO Data System.* Washington, DC: Enterprise Community Partners, Inc.

Schmitt, B. 1997. *Community Organizations and GIS Implementation: From Experience to Theory*. Piscataway, NJ: Center for Social and Community Development, Rutgers University.

Teixeira, Samantha, and John Wallace. 2013. "Data-Driven Organizing: A Community-University Partnership to Address Vacant and Abandoned Property." *Journal of Community Practice* 21: 248–62.

Theodos, Brett. 2012. Analysis of Current Population Survey, 1998–2011. Unpublished paper. Washington, DC: Urban Institute.

Theodos, Brett, Claudia J. Coulton, and Amos Budde. 2014 "Getting to Better-Performing Schools: The Role of Residential Mobility in School Attainment in Low-Income Neighborhoods." *CityScape: A Journal of Policy Development and Research* 16 (1): 61–84.

Tough, Paul. 2008. *Whatever It Takes: Geoffrey Canada's Quest to Change Harlem and America.* New York: Houghton Mifflin.

University Center for Social and Urban Research. 2011. *Foreclosure in South Pittsburgh's Hilltop and Effective Responses.* Pittsburgh, PA: University Center for Social and Urban Research, University of Pittsburgh.

US Department of Education. 2012. *Impact in Place: A Progress Report on the Department of Education's Place-Based Strategy*. Washington, DC: US Department of Education.

US Department of Housing and Urban Development. 2011a. *Choice Neighborhoods: Summer 2011 Update.* Washington, DC: US Department of Housing and Urban Development.

US Department of Housing and Urban Development. 2011b. "Choice Neighborhoods: History and Hope." In *Evidence Matters: Transforming Knowledge into Housing and Community Development Policy*. Washington, DC: Office of Policy Development and Research, US Department of Housing and Urban Development.

Von Hoffman, Alexander. 2012. "The Past, Present, and Future of Community Development in the United States." In *Investing in What Works for America's Communities: Essays on People, Places and Purpose,* edited by Nancy O. Andrews and David J. Erickson. San Francisco: Federal Reserve Bank of San Francisco and Low-Income Investment Fund.

Walker, Christopher. 2002. *Community Development Corporations and Their Changing Support Systems.* Washington, DC: Urban Institute. April.

Walker, Christopher, Sarah Rankin, and Francisca Winston. 2010. *New Approaches to Comprehensive Neighborhood Change: Replicating and Adapting LISC's Building Sustainable Communities Program.* Washington, DC: Local Initiatives Support Corporation.

Walker, Christopher, Francisca Winston, and Sarah Rankin. 2009. "Assessing Community Development Program Performance: Quantitative Assessment Strategies for the LISC Sustainable Communities Initiative." Washington, DC: Local Initiatives Support Corporation.

White House. 2011. *Building Neighborhoods of Opportunity: White House Neighborhood Revitalization Initiative Report.* Washington, DC: Neighborhood Revitalization Initiative, White House. July.

ESSAY

# Neighborhood Data and Locally Driven Community Change

Robert J. Chaskin

The past 20 years have seen a resurgence of place-based efforts to promote community change through a range of community development, organizing, and community-building strategies (Briggs 2002; Kingsley, McNeely, and Gibson 1997; Kubisch et al. 1997, 2002, 2010; Saegert 2005). From the work of individual community-based organizations (CBOs) to foundation-funded comprehensive community initiatives (CCIs) to federal policy (most recently the Promise and Choice Neighborhood initiatives), such efforts target action at the neighborhood level, seek to leverage neighborhood resources and actors as agents of change, draw on a range of strategies, and have a broad (and often ambitious) vision of neighborhood transformation that goes beyond community-based service provision or production (e.g., of housing, commercial activity, jobs) to focus as well on strengthening the leadership, social capital, and organizational infrastructure of communities and on building their capacity to manage and foster change (Chaskin et al. 2001).

One increasingly recognized requirement for improving community-building practice is better information about neighborhoods and community-change processes. Indeed, there has been mounting interest in learning from community-change efforts, bringing to bear information and analysis to inform them, increasing information access, and building local capacity to use information to strengthen the hand of neighborhood-level actors in pursuing their community-change goals

(Connell et al. 1995; Federal Reserve 2011; Kingsley 1998; Sawicki and Craig 1996). This agenda presents both opportunities and challenges. Some are conceptual, such as around key aspects of community change goals and how best to measure them. Others are more technical, such as around data access and analysis. Yet others are interactional, entailing contextual, political, and organizational influences that shape interests, priorities, expectations, and use of data in these contexts.

This essay explores some of the dynamics around using data for various purposes in community-change efforts. Drawing on several empirical examples, it interrogates some of the principal ideas, uses, tensions, and dynamics that inform data use and some key aspects of process and capacity that need to be recognized and addressed. First, it outlines some arguments for the importance of neighborhood data in locally driven community-change efforts and some of the principal uses to which data might be put. It then considers some challenges and tensions that may be encountered in seeking to harness data and analysis for these purposes. Next, it explores the uses of and dynamics around data collection and analysis in the service of community-change efforts, drawing on three brief examples of CBO and CCI action. Finally, it distills some potential implications suggested by these efforts to improve data use and analysis for community change.

## Rationale and Intent

The increasing interest in better harnessing neighborhood-level data to support neighborhood-change efforts is part of a broader return to community that has taken place over the past 25 years. Research on neighborhood effects and the problems associated with concentrated urban poverty (e.g., Jencks and Mayer 1990; Sampson, Morenoff, and Gannon-Rowley 2002; Wilson 1987), as well as research on changes in civic engagement and the nature of community in contemporary society (Putnam 2000; Wellman 1979), have been accompanied by support for a number of efforts. These efforts—large and small, funded by philanthropy and government, focused on a range of social problems and goals—have been grounded in local communities as both sites and targets of change. Community-based efforts across this spectrum treat the local community (often an urban neighborhood) as both the context for and the principle around which practice should be organized. As context, community is taken account of to make policies and programs more relevant, responsive, and effective. As organizing principle, community is seen as the unit of planning and action toward which policy is directed, for which programs are developed, around which strategic goals are established, and through which activities and services are provided (Chaskin 2002).

This dual orientation to neighborhood informs a focus on data as essential both to understanding context and to informing and assessing the effects of action. These efforts also take place within a broader embrace of what might be called the empirical imperative: the increasing push toward evidence-based policy and practice and toward data-driven planning in the public, private, and nonprofit sectors (Briggs and Rzepnicki 2004; Davies, Nutley, and Smith 2000; Grantmakers for Effective Organizations 2004; Reid 1994, 2002). Finally, these efforts are often informed by notions about the relationship between information and empowerment, about the ways in which democratizing data by increasing access to information and community actors' capacity to use it can allow communities to enhance reflective practice and to harness the power of data for their own purposes (Sawicki and Craig 1996; Schön 1983).

As the outline of these orientations already begins to suggest, the value of neighborhood data and analysis are invoked for a range of uses. They can be used as a planning tool to help clarify goals and identify needs and opportunities that can be translated into action. Neighborhood-level data have served in this capacity at least since the Progressive era (with more or less emphasis over time), when community-mapping research was conducted by settlement workers, such as those associated with Hull House, to provide demographic and social-needs assessment information that could inform both service provision and advocacy activities (Addams 1895). Neighborhood data also can be used for outreach and engagement, in which data collection serves the reciprocal purpose of collecting information and involving neighborhood residents in identifying priorities, assessing community assets, and recruiting participants in community-change activities (Kretzmann and McKnight 1993). Youth mapping projects, for example, have sought to mobilize young people to collect information about their neighborhoods in ways that provide useful information not available from existing data sources, build skills and commitment among participating youth, and provide youth employment (Kaufman 2011).

Neighborhood data can be used as a tool for assessment. This use may be oriented toward—or perceived to be oriented toward—different goals. For example, evaluation activities may be designed primarily for understanding the impact of community-change activities on some range of outcome targets; as a mechanism of accountability to ensure funding obligations and program benchmarks are being met; or as a tool for ongoing learning, providing a feedback mechanism to inform community members and participants in community-change efforts about emerging challenges, lessons, and opportunities that might suggest action toward midstream course correction. These different orientations toward assessment are, of course, not mutually exclusive, as arguments for empowerment evaluation and utilization-focused evaluation attest (Fetterman, Kaftarian, and Wandersman 1996; Patton 2008).

Neighborhood data—both on context and on the nature and effects of neighborhood-change efforts—can also be used as a tool for leveraging resources by providing a way to make the case to potential funders and partners about the nature of need, the appropriateness of investing resources, and the value of supporting particular organizations or efforts (assuming, of course, that there is a positive story to tell).

Finally, neighborhood data can be used to support social action directly, addressing issues of inequality by serving as a mechanism for helping to promote social change (Nyden and Wiewel 1992; Stoeker and Bonacich 1992). Direct support for social action is the principal rationale behind participatory action research, which, drawing in part on Freirean notions of empowerment through reflective action (praxis), seeks to shape community–researcher partnerships in the service of community-led processes of investigation, analysis, and action (Freire 1990; MacIntyre 2008; Wallerstein and Doran 2003; Whyte 1989). Through these processes, data can be harnessed to lead to collective action on the part of communities and to shape "causal stories" (Stone 1989) that support claims-making regarding the nature, causes, and locus of responsibility for addressing community needs and circumstances.

## Challenges and Tensions

These various potential uses of neighborhood data to inform and assess community-change efforts are accompanied by core challenges. One set of challenges concerns the very nature of neighborhoods and of neighborhood-change efforts. Neighborhoods are complex, open systems, variously defined and subject to myriad influences beyond those shaped by any given community-change effort. This complexity complicates efforts to outline theories of change that specify causal expectations between input and intended outcome and makes identifying comparison communities to establish a counterfactual difficult. In addition, community-change efforts themselves tend to be similarly complex, seeking to address several issue areas (social, economic, physical) across sectors (public, private, nonprofit) and at different levels (individual, organization, community). They also change over time in response to emerging challenges and opportunities, seeking to be responsive to local circumstances and local priorities—all of which makes process and organization as crucial to the understanding of the possibilities and limitations of these efforts as any measure of neighborhood change. Finally, some of the outcomes these efforts seek to effect—"strengthened community capacity, enhanced social capital, an empowered neighborhood" (Kubisch, Fulbright-Anderson, and Connell 1998, 4)—are imprecisely defined and present particular measurement problems for which there are few widely accepted solutions (Chaskin 2002; Kubisch et al. 1995, 1998; Rossi 1999).

A second set of challenges concerns the nature and availability of data about neighborhoods. Existing data are of differing quality and are collected on the basis of different units of analysis (census tracts, Zip Codes, police precincts, service catchment areas) and over different periods. Although tools and techniques for aggregating, mapping, and analyzing small-area data have improved dramatically (Coulton and Hollister 1998), much information relevant to understanding community circumstances and dynamics—resident perceptions, behaviors, and relational networks, for example—is not available through existing data sources. Importantly, it is often these more elusive dimensions, such as the level of neighborhood social capital, that are of central interest to those involved in community-building efforts (Chaskin 2002; Kubisch et al. 1998).

A third set of challenges concerns the inclination and capability of community actors—organizations and individuals—to collect and use information. Much existing data are held by actors outside the community, such as government agencies, universities, and private organizations. Gaining access to this information often requires time-consuming, sometimes difficult, and not always successful negotiation with agency personnel. In addition, community actors differ in their resources and capacity to ask researchable questions, work with existing information, and engage in various data-collection strategies to collect information not available through existing sources (Chaskin et al. 2006).

Finally, as implied by the different potential uses of neighborhood data outlined above, the potential audiences and users of data are often highly diverse, including direct participants in change efforts (community organizations, local government, residents, funders) and others not directly active but for whom data may well be relevant (residents, other organizations, potential funders, the public at large). Different users often have different interests, different levels of comfort with research and evaluation and different expectations for what they can provide, and different requirements for the nature of evidence that will be useful to their planning or convince them of findings regarding impact. Recognizing that knowledge is socially constructed—recognized, interpreted, and transformed by a range of actors for a range of purposes, often in contexts of conflict, unequal power relations, and competing interests (Chaskin 2008; Habermas 1989; Huberman 1987; Kondrat 1994)—raises questions about what might count as data, for whom, and with what relevance.

## Three Case Studies

The remainder of this essay explores some of the approaches, benefits, challenges, and dynamics of efforts to use data in the service of neighborhood change through a set of three brief case studies of CCI and CBO action. Taken together, the cases provide insights into different opportunities and challenges regarding data availability, potential uses of neighborhood data, neighborhood and organizational capacity, and community dynamics around goals, expectations, findings, and value. The first case provides an example of these dynamics in the context of initiative evaluation, the second in the context of neighborhood planning, and the third in the context of an effort to use data as a tool for mobilization and social action.

### Neighborhood Data and Program Evaluation

The discussion to follow is based on the experience of an early, multicity CCI—the Neighborhood and Family Initiative sponsored by the Ford Foundation from 1990 to 2000—and its efforts to shape evaluation activities to support both cross-site analysis to inform the broader field and local evaluations tailored to the needs of each neighborhood initiative.[1] Although presenting a highly specific set of circumstances, the case highlights some of the contextual, political (with a small "p"), and organizational influences that can shape local evaluation choices and activities and that have relevance for considering various approaches to using neighborhood data in CCIs and other community-change efforts.

The two-tier evaluation strategy was embraced in light of some of the key challenges noted above regarding the complexity, evolutionary nature, and context specificity of the initiative. Although a national evaluation was to produce a detailed implementation study of the initiative over time and synthesize findings across sites, separate local evaluations were intended to address each site's specific needs for information and assessment, guided by the specification of outcome objectives (and appropriate measures) as they developed through the strategic planning and implementation processes. This strategy led to quite different orientations to local data collection and analysis across sites and some significant shifts in direction. In one site, for example, local evaluation activities were first oriented toward clarifying outcome goals as early planning evolved, and then focused more specifically around data collection and analysis, very much connected to ongoing planning. The local evaluation team participated as part of the management team of the organization, providing information on community resources, helping staff identify measures of program outcomes, and framing questions to assist staff in the running of the organization and its programs.[2]

At another site, the initial evaluation plan was organized around three separate components, each provided by a separate researcher, focused respectively on culling administrative and census data to track neighborhood change over time, documenting programs and events to tell the story of the effort, and observing and providing feedback on the planning and implementation processes.[3] Later, the evaluation focused more on collecting data from administrative sources on specific issues of interest (e.g., school performance data) and developed additional instruments to more specifically track progress on organizational development issues and on the impact of certain projects, both measured by brief survey instruments seeking to understand the perceptions of participants (Wasson 1999). Its major focus, however, remained on process and on tracking basic organizational benchmarks (e.g., filling membership vacancies, conducting training workshops) rather than on programmatic or neighborhood outcomes.

In a third location, an evaluation team of three was also assembled, this time largely in response to local concerns regarding ethnic diversity (the neighborhood being largely African American and Caribbean–West Indian) and the desire to establish a team that the community would see as legitimate. The team was led by a professor at a local university (an African American seen to be respected in the community) who was joined by researchers from two local nonprofit research organizations. One of these organizations, which was seen as being "white," was well-established and had both the capacity to conduct a survey and specific experience in community assessment in the city. The other was a new organization based in the target neighborhood and run by West Indian researchers well-known in the Caribbean–West Indian community. The partnership eventually dissolved with the neighborhood-based organization continuing as evaluator. Beyond the neighborhood survey, evaluation activities ultimately focused on monitoring benchmarks of planned activities (e.g., develop community newsletter; work with the city to identify employment opportunities) rather than on the outcomes such activities sought to achieve. Evaluation also included open-ended interviews to glean the perceptions of participants and a small set of residents about neighborhood change (Johnson and Johnson 1996).

At the fourth site, data meant to inform the planning process was collected though asset-mapping surveys (McKnight 1987), windshield surveys of housing and neighborhood conditions, and focus groups with neighborhood youth. The early phase of the evaluation was organized around the creation of a learning community meant to foster an extended conversation among a small group of participants and community members. In this way, the evaluation sought to assess initiative progress through collecting, telling, and discussing stories of goals, actions, and effects.[4] Later, in response to requests for more concrete data, the evaluation began to focus on assisting implementing organizations to establish basic management information systems with which they (and the initiative more broadly, which sponsored these activities) could track their activities and the users of their services.[5]

These brief vignettes are too summary to give the dynamics behind these choices and changes their due, but they do suggest some ways in which a range of factors influences what is possible and what is ultimately pursued with regard to neighborhood data for local evaluation. The relative lack of focus on available data and neighborhood indicators and more general focus on output benchmarks and participant perceptions were driven by several factors. One was defensive: establishing criteria against which to evaluate progress was contentious, provoking calls for greater inclusiveness, alternative modes of inquiry, and a reluctance to collect data that focused on neighborhood-level change given the sense that moving the needle on such outcomes was unlikely in light of the scale and scope of intervention. Another factor was practical: local evaluations were seriously underfunded and were provided with little technical assistance, and there was little capacity locally to collect, manage, and analyze data within these resource constraints. Further, in light of the desire for feedback on critical process issues and the inadequacy of the national evaluation in providing it in a timely, succinct, and site-specific manner, limited resources tended to be committed to providing formative feedback on more process-oriented issues regarding planning, decision-making, and implementation challenges. A third factor concerned basic orientations toward evaluation and their position vis-à-vis other priorities: in the absence of clear expectations from the funder, participants were initially content to understand change informally, through their day-to-day interactions and observations. The sense that "we'll know it when we see it" without a structured evaluation was common, and plans to address local evaluation in early proposals were largely just pro forma acknowledgments of the funder's expectation that something would be put in place. In addition, rather than providing dedicated funding for local evaluation, resources were to be taken from the general grant provided to each neighborhood initiative. Such funds were often seen as being taken from investment in programmatic efforts. Finally, certain kinds of political considerations were at play: the desire for inclusivity and an approach to local evaluation that reflected and respected community identity and process informed the search both for researchers of color and the adoption of methods privileging reflection and narrative. And, given a long history of extractive research in and negative characterizations of poor communities and communities of color by universities and (mostly white) researchers, evaluation was initially seen by some as something to protect against rather than to engage in.

## Neighborhood Data and Planning through Community-University Partnerships

The second case focuses on an effort to establish a partnership between a CBO in each of two Chicago neighborhoods and university researchers to measure aspects of community social capital within the budget, time, and skills constraints under which CBOs often must work. The effort provides insight into the potential and limitations of particular CBO-friendly measurement tools and approaches and highlights the complexity of partnership dynamics, even under the conditions of relative trust, mutual commitment, long-term relationships, and dialogue that are often cited as key variables in successful researcher-community partnerships (Israel et al. 2003; Wilson 2004).[6]

Measurement focused on three dimensions of associational action related to social capital: collective efficacy (Sampson, Raudenbush and Earls 1997), neighborhood activism, and involvement in voluntary associations. In addition to the theoretical arguments and evidence that suggest a relationship between high levels of these constructs and certain aspects of community well-being (Putnam 1993; Sampson et al. 1997, 2002), the constructs speak to the presence of reservoirs of neighborhood strengths that can be harnessed for collective purpose and suggest potential responses CBOs might launch in light of their absence.

Data collection relied on strategic convenience sampling methods. CBO staff members' local knowledge of community characteristics and dynamics was used to select sites likely to provide access to respondents who would approximate a representative sample. Results were compared with findings from a random-sample survey collected as part of the Project on Human Development in Chicago Neighborhoods.[7] For purposes of analysis, neighborhoods were defined by aggregating the Project on Human Development's neighborhood clusters to conform to the neighborhood definitions provided by the CBOs. This aggregation permitted findings to be presented at the unit of analysis most broadly relevant to the CBOs and allowed for within-community analyses at smaller neighborhood levels.

The partnership was organized as collaborative research in the sense suggested by Nyden and Wiewel (1992); researchers and CBOs each played a role in shaping the research and in collecting, analyzing, interpreting, and using the data for their respective purposes. The division of labor settled on sought to make the best use of CBO resources and expertise while minimizing the amount of CBO staff time required and without overtaxing staff capacities. CBO staff thus consulted on the development of the survey instrument and strategies of administration, coordinated and administered data collection, and acted as consumers and interpreters of the data and evaluators of the process and products developed; researchers were responsible for design, analysis, and reporting.

Some measures and venues yielded better data than others. The collective efficacy measure proved fairly robust, and although the convenience sampling led to overestimations of the level of activism and organizational involvement, it provided fairly accurate spatial patterns of the relative levels of activism. In both cases, these findings thus generated useful discussion and provided some guidance to CBOs seeking to use this information strategically.

A number of process benefits were realized, including increasing CBO knowledge about research design and data-collection methods and trade-offs, the most effective venues to get useful information from community members, and the potential benefit of short surveys for information gathering and outreach. It also fostered learning about how CBOs and researchers might most effectively work together as partners, producers, and consumers of information. But the principal benefit was the value the data provided in informing CBO planning.

Data analysis was organized to be clearly accessible to CBO partners, and CBO staff and researchers scheduled several meetings to review and reflect on the implications of the analyses. Data were displayed in several ways, including summary bullet points, tables, bar graphs, and pie charts. Graphic displays, particularly maps, proved especially useful, and discussion around these findings led community partners to think about a range of issues that informed their planning. In some cases, findings led largely to a process of interpretation, in others to the generation of new questions, and in yet others to suggestions for action.

In some cases, findings were initially surprising and generated discussion to make sense of them, and then further discussion about neighborhood circumstances of concern that might be addressed. The unexpectedly high level of homeownership in one neighborhood, for example, was ultimately reconciled to CBO perceptions by recognizing the relationship between the geographic concentration of respondents in particular parts of the community and the nature of the housing in that location. A broader discussion ensued concerning the increasing numbers of foreclosures and the impact of escalating violent crime and gang warfare on stability in the area. This discussion helped sharpen the CBO's focus on strategic action around these issues and inform their future plans. In the other neighborhood, findings suggesting significant numbers of kinship and friendship ties and reciprocated exchange generated some surprise given, for example, the amount of crime in the neighborhood. These findings also generated a set of new questions:

combined with findings suggesting relatively low levels of trust, CBO staff wanted to know more about the nature of relationships among neighbors—perhaps through focus groups with particular subpopulations (gang members, the elderly, youth)—in order to consider ways to support positive community interactions.

Similarly, a review of the ways in which varying levels of collective efficacy were patterned in one neighborhood led first to a detailed discussion about the possible reasons for such differentiation. A neighborhood cluster in the middle of one community with a low collective efficacy score was identified as an area with large numbers of apartments, high population turnover and demographic changes, increasing numbers of foreclosures, and a diminishing local organizational infrastructure. This initial discussion led to the identification of questions for further exploration, such as the desire to map incidents of crime to see the relationship between criminal activity and areas of low collective efficacy. Finally, critiques of existing programs (such as a current gang-intervention program) and early ideas for other kinds of strategic action (away from a reliance on outside youth workers and toward back-to-basics strategies such as block-club organizing) began to be raised. In connection with these discussions, the pattern of resident involvement suggested by the mapping of neighborhood activism and association membership scores provided some thoughts about where to focus organizing activity based on where the potential for leadership seemed to be concentrated.

Data—and discussion about how to interpret the data—thus led to energized consideration of action, connecting the questions posed with the strategic agendas of the CBOs. This process was supported by structured, intentional discussion, organized around reviewing the data and posing questions regarding what the data suggested. Consistent with the literature on knowledge utilization (see Hutchinson and Huberman 1993 for a review), this result has implications for the importance of interactive approaches to the dissemination and consideration of research findings. But there are costs and challenges to this as well. The level of effort, time, resources, coordination, and attention required to garner these benefits was taxing given the significant pull of CBO core activities, emerging and shifting priorities, and the day-to-day pressures governing staff allocation choices and possibilities. Further, although the research-CBO partnerships were established within a general context of trust and a division of labor was established that used the relative interests and capacities of each partner, this balance was complicated by the timeline and accountability pressures of external funding that supported the partnership as the CBOs came up against the quotidian pressures of CBO activities and other demands on their time and resources. The combination of these factors suggests that in order to strengthen CBOs' ability to gain access to and use research for planning, fundraising, advocacy, and assessment, capacity needs to be built explicitly for this purpose beyond their core organizational structure that will give them access to data, technical assistance, and analytic support and on which they can call for their purposes and on their terms.

## Neighborhood Data as a Tool for Social Action

The final case considers the use of neighborhood data as an instrumental resource to mobilize local constituencies and shape advocacy campaigns designed to catalyze policy responses, in this case, in response to the disproportionate impact of the foreclosure crisis on a neighborhood targeted by predatory and unregulated lending. Neighborhood data here served multiple purposes: they were harnessed for situational analysis, used as a community organizing tool, and leveraged to frame and support a particular causal story that challenged existing responses to the foreclosure crisis and argued for the need—and identified the actors responsible—for enacting a different solution.

The effort was driven by the Southwest Organizing Project (SWOP), a CBO with a more than 20-year history of community organizing and advocacy in its neighborhood. SWOP, along with the Greater Southwest Community Development Corporation, which was well-established in the neighborhood, served as lead agencies for the New Communities Program (NCP), a multisite CCI led by LISC Chicago and funded by the MacArthur Foundation. As an initiative, NCP has a strong data orientation, emphasizing the value of data in informing planning and implementation (with later phases of NCP increasingly focused on data-driven planning and performance management) and supporting a major evaluation. SWOP was provided support both under NCP and separately from the MacArthur Foundation to pursue its foreclosure work. The focus on data to inform organizing around foreclosures, however, began before SWOP's involvement in NCP. The effort to address the foreclosure crisis built directly on work in the 1990s to address predatory lending practices in the neighborhood, which helped establish key alliances with influential political actors such as representatives to the state legislature (including the powerful, long-serving speaker of the Illinois House of Representatives).[8] Long before the housing crisis hit with full force beginning in 2008, SWOP became aware of high rates of foreclosures in its neighborhood and commissioned research to investigate the nature, distribution, and causes of the problem (Benefield et al. 2003). Later, it began to plot data on foreclosures block by block, generating a set of detailed maps that made painfully explicit the extent and density of foreclosures concentrated in a small geographic area. SWOP drew on a range of data sources to make its case, including data on foreclosure trends over time, on foreclosure counseling demand, and on mortgage modification applications and uptake, but the visual impact of incident data presented by the maps—red dots indicating several foreclosed homes on nearly every block—was extremely powerful. "Our maps," said a lead organizer in testimony before the US Senate, "showed an entire neighborhood drowning in a sea of red."[9]

Recognizing the value of the maps to demonstrate the dramatic scale of the issue and to argue for the inadequacy of current responses, SWOP and its allies drew on them in several ways. First, they used the maps to support local outreach and mobilization—to get community leaders on board, recruit residents to tell their stories, gather additional information through local surveys administered with community partners, and solicit volunteers for direct actions. Working initially through neighborhood churches and in partnership with member organizations (including the Greater Southwest Community Development Corporation and Neighborhood Housing Services, both with expertise in foreclosures), SWOP used the data to make evident the scope of the problem, then to mobilize allies, then to reframe the nature of the debate on foreclosures away from an individualized orientation—each foreclosure the effect of an individual homeowner's inability to meet the terms of her mortgage—toward a collective orientation—a more general crisis visited on the community because of problems in the routine provision of home mortgage loans by banks and systemic flaws in the system that regulates and safeguards these transactions.

In addition to backing community mobilization, the data supported SWOP's outreach to and engagement of influential allies beyond the neighborhood, including the state representatives with whom it had worked in the context of its antipredatory lending campaigns. The work ultimately captured the attention and engaged the support of US Senator Dick Durbin who, as chairman of the Senate Appropriations Subcommittee on Financial Services and General Government, had particular interest in the foreclosure crisis. Durbin's support lent significant weight to SWOP's cause, and he was instrumental in bringing the local issue—and the narrative of the crisis as a collective rather than an individual problem—to the national stage. In December 2008, for example, Durbin held a field hearing of the subcommittee in Chicago in which, flanked by large poster-board versions of the foreclosure maps focusing on just one Zip Code within the neighborhood, the senator declared the foreclosure crisis "a cancer or a blight that's going from home to home, neighborhood to neighborhood . . . that will really threaten us if we don't do something quickly."[10] The collective orientation to the problem and to understanding its broader impact was yet more explicitly framed by a lead SWOP organizer at a subsequent public hearing in Washington, DC:

> The foreclosure crisis has, for us, shifted from being a crisis of individual families in trouble to one of an assault on the very structure of our community. As families are forced out of their homes, key neighborhood institutions are losing the social capital needed to keep them functioning, businesses are losing critical customers, and newly vacant homes are becoming havens for gangs and drug dealers. Everybody loses.[11]

Beyond their use in such public forums, the data were used to support direct negotiations with banks targeted to seek changes in loan modification review and approval processes. Access to and influence over these targets proved difficult, however. Although SWOP did ultimately negotiate a limited response from one bank that held a large number of delinquent mortgages in the neighborhood to improve uptake of existing mortgage modification programs through concerted outreach and collaboration with the bank, the overall impact was more restricted than SWOP had hoped. Still, its campaign, backed by the strategic use of neighborhood data, helped develop, deepen, and leverage important relationships with influential political actors and brought its work to the national stage. It also laid the foundation for continuing work locally to disseminate information about, mobilize responses to, and negotiate with banks about changing bank policy to more effectively engage in loan modification programs that could have a broader collective impact on the neighborhood so badly hit by the crisis (Chaskin and Karlström 2012).

## Conclusions

As these case studies suggest, neighborhood data can play a range of critical roles connected with community-change efforts and have the potential, at least, to inform planning, provide evidence about impact, and mobilize action. Maximizing their potential requires intentionality, in particular concerning fostering capacity and supporting local efforts sensitively and responsively, and facilitating provision and dissemination of data and findings in ways that make them easily accessible and generative of discussion and consideration. Maximizing the role of neighborhood data also requires promoting clarity and transparency about the intent and benefit (as well as limitations) of data collection and analysis efforts and the trade-offs inherent in different approaches, sources, and methods. Finally, it requires explicit recognition that data and data use are neither neutral resources nor neutral processes. The reasons to use data, the choices made regarding their use, and the processes of collection, analysis, interpretation, and application are embedded in social processes, are transactional, and are informed by the values, interests, priorities, and concerns of a range of actors.

## NOTES

1. For a more detailed analysis of the two-tiered evaluation and its implications, see Chaskin (2003), on which this discussion is based.

2. L. M. Gant, unpublished local evaluation proposal, NFI, Inc., 1997.

3. OMNFI Assessment Team, Orange Mound Neighborhood and Family Initiative Assessment and Documentation work plan, unpublished, 1992.

4. The Learning Community, "Initial Report of the Learning Community," unpublished, 1993.

5. The Learning Community, revised proposal for the NFI local assessment, unpublished, 1993.

6. For a more detailed analysis of this case, see Chaskin et al. (2006), on which this discussion is based.

7. Information on the Project on Human Development in Chicago Neighborhoods can be found at www.hms.harvard.edu.

8. For a more detailed discussion of SWOP's organizing work around foreclosures in the context of NCP, see Chaskin and Karlström (2012).

9. Testimony of Katie Van Tiem before the US Senate Committee on Appropriations, Subcommittee on Financial Services and General Government, April 29, 2010.

10. Quoted in John McCarron, "Dots on a Map: 'Foreclosure Cancer!'" blog post, LISC Chicago's New Communities Program, December 8, 2008, http://newcommunities.org/news/articleDetail.asp?objectID=1308.

11. Testimony of Katie Van Tiem, April 29, 2010.

## REFERENCES


Addams, Jane. 1895. *Hull House Maps and Papers.* New York: Thomas Y. Crowell & Co.

Benefield, Nathan, Samantha Dwyer, Tonia Garnett, David Hannah, and Amy Falcone. 2003. "Southwest Chicago Foreclosure Study." Technical report 10. Chicago: Center for Urban Research and Learning. http://ecommons.luc.edu/curl_pubs/10.

Briggs, Harold, and Tina Rzepnicki, eds. 2004. *Using Evidence in Social Work Practice.* Chicago: Lyceum Books.

Briggs, Xavier de Souza. 2002. "Community Building: The New (and Old) Politics of Urban Problem Solving." Faculty research working paper series. Cambridge, MA: John F. Kennedy School of Government, Harvard University.

Chaskin, Robert J. 2002. "The Evaluation of 'Community Building': Measuring the Social Effects of Community-Based Practice." In *Assessing Outcomes in Child and Family Services: Comparative Design and Policy Issues,* edited by Anthony Maluccio, Tiziano Vechiato, and Cinzia Canali (28–47). Hawthorne, NY: Aldine de Gruyter.

———. 2003. "The Challenge of Two-Tiered Evaluation in Community Initiatives." *Journal of Community Practice* 11 (1): 61–83.

———. 2008. "Research, Dissemination and Impact: Issues, Lessons and Future Directions." In *Research for Action: Cross-National Approaches to Connecting Knowledge, Policy and Practice for Children,* edited by Robert J. Chaskin and Jona M. Rosenfeld (131–157). New York: Oxford University Press.

Chaskin, Robert J., Prudence Brown, Sudhir Venkatesh, and Avis Vidal. 2001. *Building Community Capacity.* Hawthorne, NY: Aldine de Gruyter.

Chaskin, Robert J., Robert M. Goerge, Ada Skyles, and Shannon Guiltinan. 2006. "Measuring Social Capital: An Exploration in Researcher-Community Partnership." *Journal of Community Psychology* 34 (4): 489–514.

Chaskin, Robert J., and Mikael Karlström. 2012. *Beyond the Neighborhood: Policy Engagement and Systems Change in the New Communities Program.* New York: MDRC.

Connell, James P., Anne C. Kubisch, Lisbeth B. Schorr, and Carol H. Weiss, eds. 1995. *New Approaches to Evaluating Community Initiatives, Volume 1: Concepts, Methods and Contexts.* Washington, DC: The Aspen Institute.

Coulton, Claudia, and Robinson Hollister. 1998. "Measuring Comprehensive Community Initiative Outcomes Using Data Available for Small Areas." In *New Approaches to Evaluating Community Initiatives, Volume 2: Theory, Measurement, and Analysis,* edited by Karen Fulbright-Anderson, Anne C. Kubisch, and James P. Connell (165–220). Washington, DC: The Aspen Institute.

Davies, Huw T. O., Sandra M. Nutley, and Peter C. Smith (eds.). 2000. *What Works? Evidence-Based Policy and Practice in Public Services.* Bristol, United Kingdom: The Policy Press.

Federal Reserve. 2011. *Putting Data to Work: Data-Driven Approaches to Strengthening Neighborhoods.* Washington, DC: Board of Governors of the Federal Reserve System.

Fetterman, David M., Shakeh J. Kaftarian, and Abraham Wandersman (eds.). 1996. *Empowerment Evaluation: Knowledge and Tools for Self-Assessment and Accountability.* Thousand Oaks, CA: Sage Publications.

Freire, Paulo. 1990. *Pedagogy of the Oppressed.* New York: The Continuum Publishing Company.

Grantmakers for Effective Organizations. 2004. *Leveraging What You Know: Knowledge Management Strategies for Funders.* Washington, DC: Grantmakers for Effective Organizations.

Habermas, Jürgen. 1989. *The Theory of Communicative Action.* Two volumes. Boston: Beacon Press.

Huberman, Michael. 1987. "Steps toward an Integrated Model of Research Utilization." *Knowledge: Creation, Diffusion, Utilization* 8 (4): 586–611.

Hutchinson, Janet, and Michael Huberman 1993. *Knowledge Dissemination and Use in Science and Mathematics Education: A Literature Review.* Washington, DC: National Science Foundation.

Israel, Barbara A., Amy J. Schultz, Edith A. Parker, Adam B. Becker, Alex J. Allen III, and J. Ricardo Guzman. 2003. "Critical Issues in Developing and Following Community-Based Participatory Research Principles." In *Community-Based Participatory Research for Health,* edited by Meredith Minkler and Nina Wallerstein (53–76). San Francisco: Jossey-Bass.

Jencks, Christopher, and Susan Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." In *Inner-City Poverty in the United States,* edited by Laurence Lynn and Michael McGreary (111–86). Washington, DC: National Academy Press.

Johnson, C., and E. Johnson. 1996. *Evaluation Report: UANC's Initiatives.* Hartford, CT: The Center for Urban Research, Education, and Training.

Kaufman, Lisa R. 2011. *Community Youth Mapping: A Tool for Youth Participation and Program Design.* Educational Quality Improvement Program 3: Engaging and Preparing Youth for Work, Civil Society and Family Life. Washington, DC: USAID.

Kingsley, G. Thomas. 1998. "The Information Revolution." *Community: A Journal of Community Building for Community Leaders* 1 (1): 20–25.

Kingsley, G. Thomas, Joseph B. McNeely, and James O. Gibson. 1997. *Community Building: Coming of Age.* Washington, DC: The Development Training Institute and Urban Institute.

Kondrat, Mary E. 1994. "Culture and Power in Technology Transfer: Perspectives from a Critical Sociology of Knowledge." Social Development Issues 16 (3): 45–67.

Kretzmann, John, and John McKnight. 1993. *Building Communities from the Inside Out: A Path toward Finding and Mobilizing a Community's Assets.* Evanston, IL: Neighborhood Innovations Network, Northwestern University Center for Urban Affairs and Policy Research.

Kubisch, Anne, Patricia Auspos, Prudence Brown, Robert Chaskin, Karen Fulbright-Anderson, and Ralph Hamilton. 2002. *Voices from the Field II: Reflections on Comprehensive Community Change.* Washington, DC: Aspen Institute.

Kubisch, Anne, Patricia Auspos, Prudence Brown, and Tom Dewar. 2010. *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts.* Washington, DC: Aspen Institute.

Kubisch, Anne, Prudence Brown, Robert Chaskin, Janice Hirota, Mark Joseph, and Harold Richman. 1997. *Voices from the Field: Learning from the Early Work of Comprehensive Community Initiatives.* Washington, DC: Aspen Institute.

Kubisch, Anne C., Karen Fulbright-Anderson, and James P. Connell. 1998. "Evaluating Community Initiatives: A Progress Report." In *New Approaches to Evaluating Community Initiatives, Volume 2: Theory, Measurement, and Analysis,* edited by Karen Fulbright-Anderson, Anne C. Kubisch, and J. P. Connell (1–14). Washington, DC: Aspen Institute.

Kubisch, Anne C., Carol H. Weiss, Lisbeth B. Schorr, and James P. Connell. 1995. "Introduction." In *New Approaches to Evaluating Community Initiatives,* edited by James P. Connell, Anne C. Kubisch, Lisbeth B. Schorr, and Carol H. Weiss (1–24). Washington, DC: Aspen Institute.

MacIntyre, Alice 2008. *Participatory Action Research.* Qualitative Methods Series 52. Thousand Oaks, CA: Sage Publications.

McKnight, John L. 1987. *The Future of Low-Income Neighborhoods and the People Who Reside There: A Capacity-Oriented Strategy for Neighborhood Development.* Evanston, IL: Center for Urban Affairs and Policy Research, Northwestern University.

Nyden, Phil, and Wim Wiewel. 1992. "Collaborative Research: Harnessing the Tensions between Researcher and Practitioner." *American Sociologist* Winter: 43–55.

Patton, Michael Q. 2008. *Utilization Focused Evaluation.* 4th ed. Thousand Oaks, CA: Sage Publications.

Putnam, Robert D. 1993. *Making Democracy Work: Civic Tradition in Modern Italy.* Princeton, NJ: Princeton University Press.

———. 2000. *Bowling Alone: The Collapse and Revival of American Community.* New York: Simon & Schuster.

Reid, William J. 1994. "The Empirical Practice Movement." *Social Service Review* 68 (2): 165–84.

———. 2002. "Knowledge for Direct Social Work Practice: An Analysis of Trends." *Social Service Review* 76 (1): 6–33.

Rossi, Peter H. 1999. "Evaluating Community Development Programs." In *Urban Problems and Community Development,* edited by Ronald F. Ferguson and William T. Dickens (521–67). Washington, DC: Brookings Institution Press.

Saegert, Susan. 2005. *Community Building and Civic Capacity.* Aspen Institute Roundtable on Community Change working paper. Washington, DC: Aspen Institute.

Sampson, Robert J., Jeffrey Morenoff, and Thomas Gannon-Rowley. 2002. "Assessing Neighborhood Effects: Social Processes and New Directions in Research." *Annual Review of Sociology* 28: 443–78.

Sampson, Robert J., Stephen W. Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277: 918–24.

Sawicki, David S., and William J. Craig. 1996. "The Democratization of Data: Bridging the Gap for Community Groups." *Journal of the American Planning Association* 62 (4): 512–23.

Schön, Donald. 1983. *The Reflective Practitioner: How Professionals Think in Action.* New York: Basic Books.

Stoeker, Randy, and Edna Bonacich. 1992. "Why Participatory Research?" *The American Sociologist* 23 (4): 5–13.

Stone, Deborah. 1989. "Causal Stories and the Formation of Policy Agendas." *Political Science Quarterly* 104 (2): 281–300.

Wallerstein, Nina, and Bonnie Duran. 2003. "The Conceptual, Historical, and Practice Roots of Community-Based Participatory Research and Related Participatory Traditions." In *Community-Based Participatory Research for Health,* edited by Meredith Minkler and Nina Wallerstein (27–52). San Francisco: Jossey-Bass.

Wasson, Rebecca. 1999. *Orange Mound Collaborative: Growth and Change 1999.* Memphis, TN: Wasson Associates.

Wellman, Barry. 1979. "The Community Question: The Intimate Networks of East Yorkers." *American Journal of Sociology* 84 (5): 1201–31.

Whyte, W. F. 1989. "Action Research for the Twenty-First Century: Participation, Reflection, and Practice." *American Behavioral Scientist* 32 (5): 499–623.

Wilson, David. 2004. "Key Features of Successful University–Community Partnerships." In *New Directions in Civic Engagement: University Avenue Meets Main Street* (17–23). Charlottesville, VA: Pew Partnership for Civic Change.

Wilson, William Julius. 1987. *The Truly Disadvantaged.* Chicago: University of Chicago Press.

ESSAY

# Cutting through the Fog

## *Helping Communities See a Clearer Path to Stabilization*

Lisa Nelson

*"We need to know what is transpiring in our neighborhoods in order to take action and develop solutions," notes Frank Ford, formerly of Cleveland Neighborhood Progress, a community development corporation intermediary in Cleveland, Ohio. "Data help pull back the curtain—clear away the fog—identify what is happening and help guide our efforts going forward."*

Data are integral to many neighborhood initiatives throughout the City of Cleveland, Ohio, and its county, Cuyahoga County. At the center of these initiatives is Northeast Ohio Community and Neighborhood Data for Organizing (NEO CANDO), a data system developed and maintained by the Center on Urban Poverty and Community Development at Case Western Reserve University in Cleveland. Launched in 1992, CANDO, as it was previously called, evolved from a system that required users to dial in by modem and contained mainly social and economic indicators aggregated to various geographic levels. Renamed NEO CANDO in 2005, the system moved to the web and expanded to include data from other counties in northeast Ohio. Today's system allows web-based, easily searchable access to real property information at the parcel level and in real time. The NEO CANDO system, which continues to evolve and expand both technologically and in data holdings, moved from a useful tool in decisionmaking to an essential one. This essay describes the development of the data system and discusses various ways the data and partnerships are leveraged to inform neighborhood improvement initiatives.

## Early Days

Back in 2004, Neighborhood Progress, Inc. (NPI), a community development corporation (CDC) intermediary in Cleveland, along with local community development organizations, undertook a strategy called the strategic investment initiative (SII).[1] SII was designed to target investments in six Cleveland neighborhoods with the goal of market recovery and stabilization. Neighborhoods chosen for funding under this initiative were those with community assets that could be built upon—strong anchor institutions, organizational capacity, and indications of market strength or potential market recovery. Median sales prices, vacancy rates, and homeownership rates from NEO CANDO were used to identify signs of market strength. At the time NEO CANDO offered limited parcel-level data—namely sales transactions and property characteristics from the Cuyahoga County auditor—that were important but not sufficient to assess the status of properties in these neighborhoods. Instead, parcel-level data on property condition and vacancy status were available via a Palm Pilot survey that had been administered in the six SII neighborhoods. But the organizations involved in SII knew they lacked other key pieces of information about properties, such as code violations or whether the property was in foreclosure. These missing pieces could be cobbled together with time and effort, but what became clear was the need for this information to be integrated into one data system, eliminating the time-consuming and cumbersome processes of retrieving data from various data systems across county agencies. Having such a system would require less time for acquiring data and allow more time for making decisions about what to do with properties. With NEO CANDO, the infrastructure necessary for a one-stop shop for parcel-level data was in place, as was staff capacity to make it happen.

During the same time, many of Cleveland's neighborhoods were plagued with increasing numbers of vacant and abandoned properties. Concerns were growing about the negative effects these properties would have on the progress made in revitalizing and strengthening neighborhoods over the previous decade. To address this problem, the National Vacant Properties Campaign assessed the vacant and abandoned properties in Cleveland and recommended strategies to prevent, reclaim, and reuse these properties. In their summer 2005 report, the authors wrote that the lack of comprehensive information about real property in the city contributed to the lack of progress in addressing vacancies (Mallach, Mueller Levy, and Schilling 2005). The authors argued that if city decisionmakers didn't know how many vacant properties there were, where they were, and their impacts on neighborhoods, then developing solutions would be difficult. The National Vacant Properties Campaign recommendations laid the foundation for a strong push by the broader Cleveland CDC community for access to property-level data on code violations, condemnations, and demolitions as a way to better identify these problem properties and begin to counteract their negative impacts.

Leading the way in this effort was NPI, which convened key stakeholders and decisionmakers across the county to push for access to property-level data and develop strategies to deal with vacant property issues. The Vacant and Abandoned Property Action Council (VAPAC), as the group became known, included community development organizations that would use the data; city and county agencies that collected the data; Case Western Reserve University's NEO CANDO staff, who had the technical expertise to make the data accessible; and finally, the funders who were being asked to support the upgrade of NEO CANDO to house these data. VAPAC, in its early meetings, pushed to have decisionmakers at the table—in other words, individuals who had the authority and ability to provide data. The group began meeting regularly in late 2005 to determine which data were critical to acquire, develop strategies to acquire the data, and discuss how to make them available through NEO CANDO.

Much of the data VAPAC needed were stored in Cleveland's Building and Housing Department, which collects and maintains information on code violations, condemnations, building permits, and demolitions. Another important source of data was the Water Department, given that water shut-off can signal vacancy. Likewise, as a foreclosure filing can be a precursor to vacancy, knowing that a property is being foreclosed on is essential to intervening with the borrower and staving off another vacant home.

VAPAC's efforts were timely. With the SII project under way and the city's glut of vacant properties and those at risk for vacancy on the rise, the group had to get a better handle on the status of these properties. However, obtaining the data was not an easy task. In some cases, the data were not stored electronically. In other cases, data providers expressed concerns about the completeness of the data. Some data providers did not see the importance or value of making these data available, nor was doing so a priority. In other instances there were actual or perceived technical issues in transferring data from one format to another. Although the VAPAC meetings could get uncomfortable at times, all parties involved were committed to and focused on developing strategies to stabilize Cleveland's neighborhoods—and data were essential to that effort.

Fast-forward to 2010 and the launch of the Neighborhood Stabilization Team (NST) web app, a component of NEO CANDO, which contains regularly updated property-related data for Cleveland and allows the user to filter the data based on geography or other characteristics, download the data, and map them. Included in this system are the very sources of data discussed in VAPAC meetings from years ago, such as code violations, permits, condemnations, and demolitions. Since 2010, these data have been made available weekly by the City of Cleveland (Hirsh, Schramm, and Coulton 2012). Access to these city data sources is limited to select users such as CDCs and municipal governments, but other parcel-level data such as sales transfers, foreclosure filings, and sheriff's sales, which were previously available only monthly or annually, are now available weekly through the publicly accessible NEO CANDO. Through the determination and commitment of those individuals sitting around the table at those early VAPAC meetings, crucial and timely data about the status of properties are now available and accessible to those seeking to bring about change in their neighborhoods.

## Neighborhood Stabilization Team and Neighborhood Stabilization Program

The SII project team was at the forefront of using parcel-level data to take action in Cleveland neighborhoods. They began meeting monthly in 2005 to identify properties for possible acquisition in the six neighborhoods where these targeted investments were being made. Staff from NPI, the neighborhood's community development organizations, NEO CANDO, and Cleveland State University's Marshall College of Law would sit in a room with a large map of an SII area. A detailed legend explained the map's colors, symbols, and patterns, which depicted various attributes (including condition, vacancy status, and tax delinquency status) of the properties within the SII area. By looking at the map, the team could see immediately the condition of the property and whether the property was tax delinquent or vacant or both. Armed with this information, the team would make strategic decisions about specific properties in the area, particularly whether a vacant property should be acquired and, if so, whether its condition indicated that it should be renovated or demolished. The map also revealed distressed (tax-delinquent or vacant) properties located adjacent to properties where investments had already been made. The team used this information to decide whether to acquire such properties to eliminate the negative impact they might have on already-acquired properties.

As the foreclosure crisis hit the Cleveland area with brute force, the number of vacant and abandoned houses that neighborhoods were already dealing with grew larger. In response to the crisis, the federal government created the Neighborhood Stabilization Program (NSP) to help communities mitigate the negative impact of foreclosures and resulting vacancies. The first round of NSP funds was awarded to individual cities or counties. Round two, however, required more collaborative efforts. The Cleveland–Cuyahoga Consortium for NSP2 was formed with the Cuyahoga Land Bank as the lead and Cuyahoga County's Development Department, the Cleveland Department of Community Development, the Cuyahoga Metropolitan Housing Authority, and several nonprofit organizations as partners. These organizations came together and developed a proposal for NSP2 funds that built on the data-driven model used in the six SII neighborhoods (Cleveland–Cuyahoga Consortium for NSP2 2009). Areas identified for NSP2 funding were selected on the basis of need and opportunity. Need was determined by using comprehensive data on home sales activity and residential vacancies from NEO CANDO, and opportunity was based on an area's community assets and organizational capacity. Of the 20 target areas chosen for this funding, 15 were in the city of Cleveland and 5 were in the inner-ring suburbs of the county.

NST worked closely with the 20 CDC and municipal groups who worked in the areas targeted for NSP2 funding. Central to the NST approach to stabilization was using up-to-date parcel-level data. Just as critical, NST's approach involved engaging with people and assisting them in the best strategic use of the data to make decisions about problem properties. The team's strength lay in its collective expertise: it included individuals with data and legal expertise as well as individuals with on-the-ground knowledge about what was happening in neighborhoods. As they did earlier in the SII, members of NST came to monthly meetings with color-coded maps showing the status of the properties in distress (vacancy, foreclosure, or poor condition) within a targeted area. At these meetings, team members immediately accessed up-to-date information about these properties (including liens, code violations, property values, and property size) and determined whether specific properties should be acquired and rehabbed or whether they should be demolished.[2]

Preventing foreclosures is one way to lessen the number of homes entering vacancy and contributing to neighborhood blight. In partnership with Empowering and Strengthening Ohio's People (ESOP), a foreclosure prevention agency, NST used foreclosure filing data (available from the NEO CANDO system from early 2007) to generate lists of borrowers whose homes had entered foreclosure, and ESOP used the lists to reach out to these borrowers with hopes of helping the owners keep their homes. Similarly, data on adjustable-rate mortgages, linked with mortgage deeds, were used to identify individuals at risk for foreclosure. ESOP staff also reached out to these borrowers in hopes of keeping them from entering the foreclosure pipeline.

NST and its community partners wanted to protect areas targeted for NSP2 funds and other public and private investments from further deterioration and instability. This interactive work between the NST and CDC and municipal staffs allowed for data-based decisions. For example, a property entering foreclosure located right next to a recently renovated property might be a prime candidate for door-knocking by ESOP. If code-violation data revealed the need for a major property repair, the homeowner could be contacted and provided assistance to mitigate the risk of further property deterioration. Having the ability to immediately identify problem properties with a few keystrokes allowed for swift and immediate action to be taken to protect long-term investments in these communities. Time was spent identifying and addressing destabilizing factors rather than searching for data.

Again and again, organization and municipal groups reported the value of having access to timely and frequently updated data in a single system and the ease with which data could be extracted from it. Greg Baron, housing director at the Stockyard, Clark-Fulton and Brooklyn Centre Community Development office, spoke of using the NEO CANDO system to find properties for the office's single-family rehab program.[3] The Cuyahoga Land Bank acquires lots of vacant and abandoned properties in its inventory. The users of the NEO CANDO–NST web app, like Baron, can pull up properties in their service area, identify the ones held by the land bank, immediately access the property characteristics of these properties, see the status of the other properties on the street, and determine whether a particular property is one their organization wants to invest in.

Another useful feature of the NEO CANDO–NST web app, according to users, is the ability to upload one's own data into the system. Baron uses this system as his primary database for properties in the Stockyard, Clark-Fulton and Brooklyn Centre's service area. According to Baron, it made sense to add the organization's data to the NEO CANDO system rather than creating a separate database. The CDC administers a vacancy survey twice a year in their neighborhood. Once the survey is complete, they upload the data into the NEO CANDO–NST web app, which contains two other sources of data on vacancy: one from the US Postal Service and another from the City of Cleveland's vacancy survey. Having three sources of vacancy data gives a better idea of a neighborhood's vacancy rate, and having a property verified by all three sources provides more confidence that it is actually vacant. Simply knowing whether a property is vacant, along with information about liens and code violations, is important in deciding whether to invest time in pursuing the property for possible acquisition. "The timesaving aspect of having the relevant property information in one system cannot be overstated," claims Baron. "It saves staff time and increases work efficiency."

The NEO CANDO system is also used for advocacy and community outreach. NEO CANDO's geographic information system was used by the Stockyard, Clark-Fulton and Brooklyn Centre CDC to create neighborhood block club boundaries. Once uploaded to the NEO CANDO system, these boundaries enabled them to generate reports of distressed properties for concerned residents. Let's say block club members in one west-side Cleveland neighborhood are interested in condemned properties on their streets. Within minutes, the CDC staff can choose a specific block club area from a drop-down menu, extract all properties that have been condemned by the city in the past month, and prepare a list for the block club members showing who owns the property, when the property was condemned, and the status of the condemnation. With this information, block members keep an eye on these properties, track whether steps are being taken to deal with these public nuisances, and reach out to public officials to deal with them in a timely manner.

Community organizers within the CDC can also generate lists of homeowners having financial difficulty with their mortgage or taxes and reach out to them through a door-knocking effort. When a property is in tax foreclosure and vacant, the CDC can work with the county to fast-track the foreclosure through the county's Board of Revision, which allows the property to be moved into the hands of a fiscally responsible owner such as the land bank or the CDC more quickly. Identifying these properties in a timely manner allows action to be taken more swiftly, which in turn decreases the amount of time a property sits in disrepair and increases neighborhood blight.

Another user of the NEO CANDO–NST web app is the Reimagining Cleveland project, which received NSP2 funding to make use of the numerous vacant side lots throughout the city. This project allowed homeowners living next to a vacant lot the opportunity to acquire it for a community garden, a child's play area, or just to expand their yards.[4] Using the NST web app, Reimagining Cleveland was able to identify, among over 30,000 vacant parcels, those parcels that met the eligibility criteria of the NSP2 program. Eligible lots had to be located in areas targeted by NSP2 or SII or in other areas where targeted investments had taken place. Eligible homeowners could not be tax delinquent, have any code violations, or be in foreclosure. Having the NST web app allowed program administrators to quickly identify candidates for this side lot program and reach out to them to gauge interest. The versatile web app also collected and tracked homeowner applications for these side lots.

## Real Estate–Owned Property Disposition

NEO CANDO–NST web app data have also been used to investigate real estate–owned (REO) disposition practices. In one well-known example, data from NEO CANDO were included in a lawsuit filed against Wells Fargo and Deutsche Banks in 2008.[5] Using the data system, NST identified these banks as holding a large number of Cleveland's REO properties. Digging deeper into these transactions, the team found that many REO properties were tax-delinquent and that the two banks were transferring them in bulk to investors without having paid the taxes. On top of that, once transferred, many of these properties were in such bad shape they had to be demolished. Basically, the banks were buying and selling properties that were unfit for consumption.

The suit brought against these banks was to deem these properties a "public nuisance"—unsafe, uninhabitable—and to prohibit the transfer of properties until problems were abated at the banks' expense. The suit focused specifically on the problem REO properties in NPI's SII areas given the long-term investments in these areas. Although the lawsuit was not a success in the courtroom, these banks did voluntarily demolish some properties in these areas. The lawsuit also put banks on notice that the community was paying attention and would continue to push for practices that counteracted, not contributed, to blight.

REO disposition practices were also the focus of often-cited research conducted by the Center on Urban Poverty and Community Development. One of their analyses showed that upward of 65 percent of the REO properties in Cleveland were being sold out of REO for $10,000 or less in 2008, up from less than 10 percent in 2005 (Coulton, Schramm, and Hirsh 2008). Properties sold subsequent to being purchased out of REO would often become tax delinquent, and if they sold again, the sale occurred quickly, which signaled that very little property improvement had taken place (Coulton et al. 2008).

Concerns were mounting about the increasing numbers of REO, vacant, and abandoned properties, the condition of these properties, and their impact on communities. By 2007, a county land bank was being discussed as one solution. In fact, findings from the Center on Urban Poverty and Community Development were used as evidence at a state legislative hearing regarding the need for a land bank. Without access to NEO CANDO's parcel-level data on sales transactions, which included buyer and seller names and sales prices, it would not have been possible to document and quantify what was transpiring in these communities.

## Strategic Decisionmaking at the Land Bank

The Cuyahoga Lank Bank began operations in summer 2009. According to Mike Schramm, the land bank's director of information technology and research, the data in NEO CANDO are essential to the workings of the land bank.[6] Rather than replicating efforts, the bank data system connects directly with the NEO CANDO–NST web app to capture the necessary property information. Having acquired over 2,000 properties since its inception, the land bank must know the characteristics and location of properties in its inventory as well as those of the surrounding properties. Also, with the volume of land bank–eligible properties and limited resources, the land bank needs to be able to make quick, strategic decisions about which additional properties to acquire. NEO CANDO allows it to do precisely that.

The mission of the land bank begins with the words "strategically acquire properties" (Keating 2011, 8). To that end, the NST web app has been essential. Using the data and functionality in the NST web app, the land bank has developed several analytic tools that help make its acquisition and demolition decisions more strategic and efficient. Although these tools are useful citywide, they also enable decisionmakers to identify properties within designated investment areas. One tool called "the Eye" works like this: A parcel in the land bank (or one that may be acquired) is entered into the data system. Concentric circles, using spatial techniques, are drawn around this parcel to show what properties surround it. Specific property information on each neighboring property within these circles—from tax delinquency to vacancy status—can then be viewed. With this information, the land bank can make decisions about acquisition of a property and whether any surrounding properties would be candidates for acquisition. If, for example, the surrounding properties are in close proximity or adjacent to homes where investments have been made or where community assets exist, the land bank will use this information in its acquisition decisionmaking.

Land bank officials also use the Eye to spatially assemble parcels to determine whether there is enough acreage for a larger-scale development, such as an urban farm or open green space, in a community. With the Eye it is possible to determine feasibility, but the data in the NEO CANDO–NST web app are crucial for identifying properties as candidates for assembling. As a decisionmaking strategy, the Cuyahoga Land Bank classifies properties based on whether properties are under its control (properties acquired through tax foreclosure or donations from banks or governmental agencies) or might be in the future. Candidates for future acquisition might include tax delinquent properties, REOs, and vacant lots, all of which can be identified through NEO CANDO–NST web app data. Using the analytics of the Eye and data from NEO CANDO, properties are identified, the acreage of parcels is summed, and a report is generated showing land available for large-scale assembly. The overall goal is to make wise decisions about property acquisition with a focus on eliminating blight and returning land and properties to productive uses that contribute to stability rather than detract from it.

Foreclosures on vacant, tax-delinquent properties are a primary method of property acquisition for the Cuyahoga Land Bank. However, the county can foreclose on only about one-quarter of these properties in a year due to budget constraints. Consequently, the county must decide which vacant, tax-delinquent properties make the most sense to send through the foreclosure process. Currently, the decisions are based on length in delinquency and amount of delinquency. However, properties that go through the tax foreclosure process but are not wanted by the community end up on the state forfeiture list and continue to deteriorate. Working closely with the Cuyahoga County treasurer, the land bank proposes a more strategic approach. Priorities for the tax for closure pipeline and ultimately the land bank would be those properties in targeted investment areas, such as the SII and NSP2 areas; those with a specific end use; and those smaller than 4,000 square feet. By using the NEO CANDO system, the land bank can filter the list of vacant, tax-delinquent properties eligible for tax foreclosure by investment area, proximity to redevelopment projects, and square footage to narrow the list for tax foreclosure. By being more strategic and intentional about what properties should go through the tax foreclosure process, limited resources can be used more wisely.

## Code Enforcement Partnership

The NST web app is also used by CDCs as part of the Code Enforcement Partnership. In place since 2011, the partnership includes Cleveland's Building and Housing Department, city council members, and the CDCs. The partnership works like this: All code violation complaints are funneled through the city's data system; routine complaints are geocoded with parcel and ward and are then provided to the appropriate CDC in each city ward.[7] Concentrated inspection areas (CIA) are used to prioritize where initial inspections will be focused. There are about 18 CIAs in a ward and, on average, about 300 homes in a CIA. The entire ward is inspected over a six-year period; however, with limited resources, the CDC and councilperson determine the order in which CIAs will be inspected. Decisions about where to focus concentrated inspections are based on data found in the NST web app, such as the foreclosure and vacancy data and the location of targeted investments.

The routine complaints that come into the city's complaint center are uploaded into the NST web app and assigned a CIA. A CDC staff member can then identify all necessary information about the property, including who owns it. A list can be generated of those properties with a complaint; a CDC staff member inspects the home named in the complaint, confirms there is a violation, and contacts the homeowner to get voluntary compliance. This process provides an opportunity to reach out to residents and make them aware of the home-repair assistance programs available in the city, with the ultimate goal of improving property conditions and neighborhood stability.

If the violation is not taken care of, the CDC can notify city inspectors, who will then inspect the property. But here the CDC must prioritize as well, identifying those properties in the worst shape or those most problematic for a neighborhood and directing inspectors to pursue those first. Using the NST web app, the CDCs can map properties with violations and determine their locations relative to community assets or redevelopment projects. The NST web app is also used by some CDCs to enter and maintain information on their routine inspections alongside information uploaded from the city system into NST on actions taken by inspectors. Before this partnership, there was no strategy for which complaints should take precedence, nor was there the ability to track the status of complaints. The existence of the NEO CANDO system facilitates this more strategic approach to code enforcement.

## Tax-Lien Certificate Sales

NEO CANDO has served as a powerful force in bringing together communities to address shared challenges and come up with collective, more regional solutions. VAPAC, which formed in 2005 to assess and address the problems of vacancy in Cleveland, now includes representatives from Cuyahoga County's inner-ring suburbs. In its early years, VAPAC focused on data acquisition for property remediation; now VAPAC members use the data not only to document and address issues, but also to push for changes in policies that will facilitate more strategic property disposition strategies.

Here's one example. Cuyahoga County sells tax-lien certificates to investment firms on properties with delinquent taxes. In turn, these firms attempt to collect back taxes. In most instances property owners pay up. In the case of tax-delinquent properties that are vacant or condemned and whose owners cannot be located, the investment firms lose interest in pursuing payment because the likelihood of collecting is extremely low.[8] As a result, these properties continue to sit vacant, further deteriorating and contributing to community blight. To counteract this problem, staff from NPI and VAPAC began working with the county's fiscal office to secure the list of tax-delinquent properties before lien certificates for them were sold. After the list is uploaded to NEO CANDO, any property flagged as vacant or condemned or of interest to the land bank, a CDC, or a municipality is excluded from the group of tax liens sold to the investment firms. These exclusions allow the joined groups an opportunity to get vulnerable properties into the hands of those who will take actions on the property that reduce further deterioration, both to the home itself and to the immediate neighborhood. Without the up-to-date property information housed in one location, it would be difficult to identify whether the tax-delinquent property is vacant or has been condemned. Also, geographic identifiers in the system are essential in tagging the property's location, which can then be shared with the appropriate CDC or municipality or the land bank if the property is in a location of interest.

Kamla Lewis, the director of neighborhood revitalization for Shaker Heights and a VAPAC member, credits NEO CANDO with promoting a more regional approach to addressing fallout from the housing crisis. Having a data system such as NEO CANDO in place has "allowed us to see our shared similarities—not just our differences—which is the fundamental basis of building partnerships and developing solutions."

## Conclusion

By design, the NEO CANDO system puts data into the hands of those who can take action. As the system has evolved over the past 20 years, it has become essential to the work of many organizations throughout the county. It has been used to identify emerging issues, document trends, inform decisionmaking, and facilitate partnerships. The interaction between the users, developers, data providers, and funders of the NEO CANDO–NST web app has helped prioritize data needs, improve system usability, and increase the understanding of and ability to address challenges within neighborhoods throughout the county. As a member of VAPAC and longtime supporter of NEO CANDO said, "Today, we have what we need. Without reliable, up-to-date information, we are walking in the dark."

## NOTES

1. In 2013, Neighborhood Progress Inc. merged with two other organizations, the Cleveland Neighborhood Development Coalition and LiveCLEVELAND. The combined organization is now named Cleveland Progress.

2. Justin Fleming (program officer, Cleveland Neighborhood Progress), interview in Cleveland with the author, September 13, 2012.

3. Greg Baron (housing director, Stockyard, Clark-Fulton and Brooklyn Centre Community Development Office), phone interview with the author, September 25, 2012.

4. Justin Fleming, interview.

5. Frank Ford (formerly vice president for research and development, Cleveland Neighborhood Progress), interview with the author, September 13, 2012.

6. Michael Schramm (director of information technology and research, Cuyahoga Land Bank), interview with the author, September 18, 2012.

7. Philip Star (executive in residence, Cleveland State University, Maxine Goodman Levin College of Urban Affairs, and consultant for the Code Enforcement Partnership in Cleveland), interview with the author, October 13, 2012.

8. Justin Fleming of Cleveland Neighborhood Progress, and Kamla Lewis of Shaker Heights both provided information on the tax-lien sale process.

## REFERENCES

Cleveland Cuyahoga Consortium for NSP2. 2009. Neighborhood Stabilization Program 2 Application. Cleveland, OH: Cuyahoga County Land Reutilization Corporation. http://www.cuyahogalandbank.org/documents/NSP2_Application_cclrc_web_posting.pdf

Coulton, Claudia, Michael Schramm, and April Hirsh. 2008. "Beyond REO: Property Transfers at Extremely Distressed Prices in Cuyahoga County, 2005–2008." Cleveland, OH: Center on Urban Poverty and Community Development, Case Western Reserve University. http://blog.case.edu/msass/2008/12/09/20081209_beyond_reo_final.pdf.

Hirsh, April, Michael Schramm, and Claudia Coulton. 2012. "Neighborhood Stabilization Team Web Application." Research Summary 12-04. Cleveland, OH: Center on Urban Poverty and Community Development, Case Western Reserve University. http://blog.case.edu/msass/2012/09/13/Briefly_Stated_No_12-04_NST_Web_App.pdf.

Keating, W. Dennis. 2011. "Cuyahoga County Land Reutilization Corporation: The Beginning, the Present and Beyond, 2009–2011." Cleveland, OH: Maxine Goodman Levin College of Urban Affairs, Cleveland State University. http://www.cuyahogalandbank.org/articles/CCLRC_2009_2011_Report.pdf.

Mallach, Alan, Lisa Mueller Levy, and Joseph Schilling. 2005. *Cleveland at a Crossroads: Turning Abandonment into Opportunity*. Cleveland, OH: National Vacant Properties Campaign. http://www.clevelandhousingcourt.org/pdf/at_the_crossroads.pdf.

# 6

# Using Data for City and Regional Strategies

In chapter 5, we focused on how neighborhood data can be used within neighborhoods to promote development and change. In this chapter we turn to the use of data about neighborhoods to support action at other levels, such as the city, region, state, or nation. In this way, policy and program decisions that are made in those jurisdictions can be informed by data that reflect variation among neighborhoods. Neighborhoods and their particular history, character, needs, and assets can be taken into account more effectively and equitably when decisions are based on data rather than impressions or limited experience. Armed with data about their neighborhoods, advocates can collectively raise awareness of their concerns with government and civic leadership at various levels. Moreover, cities and regions can be more effective when they have the capacity to scan data on conditions and quality of life in all neighborhoods, as these are the building blocks of successful societies and competitive regional economies.

In this chapter, we provide a range of exemplary types of applications in which well-formulated neighborhood-level data and analysis have had important effects on decisionmaking and outcomes in various jurisdictions. The examples demonstrate the range of stakeholders that uses cross-neighborhood data, from neighborhood residents and institutions working to demand better city services to advocates collaborating to influence state policy. The case illustrations also represent an array of purposes to which neighborhood data are applied, ranging from enabling agencies to target scarce resources to raising public awareness about emerging social or health concerns to informing the design of interventions to address those problems. Undergirding all the examples are two basic principles: the importance of bringing together data across silos and the central role of partnerships and community engagement.[1]

## The Need for Neighborhood Data to Inform Larger Jurisdictions

The ability to compare data across neighborhoods is important for several reasons. First, there is evidence of considerable variation in neighborhood conditions and population well-being depending on place. National, state, or even municipal measures of social indicators mask these differences. Indicators for particular neighborhoods can reveal populations in great distress even when things are generally improving for people overall. Indeed, the existence of problems in health, economic opportunity, or service quality is often starkly revealed when indicators in one neighborhood are compared with another or with regional averages. Such differences often provide justification for changes in public policy, program delivery, or distribution of resources, as demonstrated by the three in-depth case studies in this chapter. Disparities in human well-being can guide efforts to mobilize communities to act on improving conditions, and neighborhood indicators that reveal pockets of concern can be used to target resources to areas where they are needed most.

Second, the growing recognition that place matters has increased the awareness of the need for neighborhood indicators in research and practice. The consensus about the powerful influence of place has been driven in large part by the virtual explosion of the scientific literature on place-based disparities and neighborhood effects on the life course of individuals (van Ham, Manley, Bailey, Simpson, and Maclennan 2012). Many studies demonstrate that disadvantaged neighborhoods have higher rates of negative outcomes for children, youth, and adults that are of societal concern, such as poor school performance, antisocial behavior, health problems, and victimization (Ellen and Turner 1997; Leventhal and Brooks-Gunn 2000). Place-based inequality in access to decent housing (Mueller and Tighe 2007), good schools (Duncan and Murnane 2011; Nelson and Sheridan 2009), job opportunities (Fernandez and Su 2004), and transportation options and healthy environments (Weiss et al. 2011) are also well-documented. Chapter 7 explores the methodological issues of understanding neighborhood effects on individuals and households. Although the mechanisms that are responsible for place-based disparities in outcomes and access are the subject of research and scientific debate (Friedrichs, Galster, and Musterd 2003; Sampson, Morenoff, and Gannon-Rowley 2002; Shinn and Toohey 2003), indicators of well-being cannot be assumed to be uniform and are likely to vary systematically within cities and regions. This reality makes data on neighborhoods vital to guide effective action on many fronts.

Third, social justice concerns have evolved to incorporate injustices attached to where people live in addition to analyses based on race, class, gender, or other attributes. By looking at disadvantaged neighborhoods through an equity lens, the focus shifts from individual deficits to an understanding of fundamental determinants of disparities in human well-being. Many of these determinants, such as access to healthy foods or exposure to interpersonal violence, vary by neighborhood and call for solutions that rectify the unequal distribution of resources within cities and regions. Neighborhood data are required both to document and reduce these place-based disparities, as illustrated by the neighborhood walk scores and the limited supermarket access area scores described in chapter 3. As another example, the Kirwan Institute at Ohio State University[2] has worked with a number of cities to prepare "opportunity maps" that use neighborhood data to visualize the clustering of multiple structural impediments to opportunity (e.g., low-performing schools, inadequate housing, lack of good jobs). Data on place-based disparities, combined with scientific evidence that inequality within metro areas is an impediment to growth (Benner and Pastor 2012), is becoming a potent force in advocating for policies that address these spatio-structural elements of inequality. The Sustainable Communities Grant Program described below reflects this growing awareness of the connections between a thriving economy and equitable development.

## Putting Neighborhood Data to Use in Cities and Regions

The successful application of neighborhood data to city or regional concerns requires more than data. It has to start with an awareness of the emerging issues and conversations that are taking place and where neighborhood data can advance the process. This awareness requires knowing the potential audiences for the information and being positioned to provide the data in a timely fashion when they can make a difference. Chapter 2 describes how NNIP partners are embedded in their communities and are entrepreneurial in this regard, identifying issues for which data would enhance the policy debate or decisionmaking. By virtue of their knowledge of the neighborhoods and data availability, they are often at the table and able to steer the discussion toward the role that data can play in formulating strategies and solutions. Foremost on their agenda is assuring that the policy or issue under discussion is informed by an appreciation of the variation among neighborhoods and the values of community engagement and equity.

Local actors organize their efforts for community change in different ways. The next two sections review common approaches that rely on neighborhood-level indicators to motivate stakeholders, target resources, and inform policy and program responses. The following section offers three case studies from Austin and Dallas, Texas, and Camden, New Jersey, that all involve using neighborhood data to address local concerns, describing the players, strategies, and results to date.

## Community Indicators Projects

Community indicators projects represent one approach for using neighborhood indicators to motivate collective priorities and action. Organizers of community indicators projects select indicators that relate to a set of community goals. These projects may span topical domains with a lens of measuring quality of life or sustainability. Others may focus on assessing the well-being of a particular group, like children or the elderly, or on a sector, such as health or arts and culture.

The selection process may be conducted in a relatively short amount of time by a small advisory committee or be designed to engage broader segments of the community as well as subject matter experts. The projects aim to update the indicators recurrently (most often annually or biannually) and sponsor a periodic review by stakeholders on whether things are getting better or worse.

The projects vary on how closely they are linked to policy or program decisions. The design of the review process may incorporate developing action plans to "move the needle" on certain indicators. Other projects may serve to raise awareness of conditions and trends generally, but in the interest of being neutral may not explicitly tie indicators to policy recommendations.

Although community indicators projects are emerging across the country, only a few drill down to the neighborhood level. The number of projects that include subcity indicators has risen as a result of the technological and organizational advances described in earlier chapters. Several of the most prominent examples are hosted by National Neighborhood Indicators partners (see table 6.1). This chapter's second case study describes in more detail the role of the Wholeness Index in raising the profile of inequities between North and South Dallas.

Merrick and Martin's essay accompanying this chapter describes the Regional Equity Atlas sponsored by the Coalition for a Livable Future and developed by the Institute for Portland Metropolitan Studies at Portland State University in Portland, Oregon. The Atlas reinforced the inclusion of social equity in the regional policy conversation around access to open space. The Coalition for a Livable Future's neighborhood data and maps supported the advocacy for a new vehicle for funding to address the inequities that the Atlas revealed. In a separate examination, Merrick contrasts the origins and design of the Equity Atlas to the later-developed Greater Portland Pulse, concluding that the geographic specificity and explicit frame of equity facilitated the use of the Atlas data in influencing policy and targeting resources.[3]

The Boston Indicators Project, implemented by the Boston Foundation and the Metropolitan Area Planning Council, has provided a model for many other local projects. The project's organizers recognize the

**Table 6.1.** Selected Community Indicators Projects with Neighborhood-Level Data

| Location | Indicators Project Name | Years |
|---|---|---|
| King County, WA | Communities Count | 1997–present |
| Baltimore, MD | Vital Signs | 2000–present |
| Greater Philadelphia, PA area | Where We Stand | 2004–2010 |
| Boston, MA region | Boston Indicators Project | 2004–present |
| Dallas, TX | Wholeness Index | 2006–2008 |
| Portland, OR-WA metropolitan area | Regional Equity Atlas | 2007–present |
| Pinellas County, FL | Pinellas Indicators | 2011–present |
| Greater Portland, OR area | Greater Portland Pulse | 2011–present |

challenge of getting residents and civic leaders to draw insights from the data that will be useful in decisionmaking. To facilitate this engagement, the project features broad-based civic gatherings to "elicit views on key long-term trends, major developments and accomplishments, and key challenges in different topic areas" (General Accountability Office 2011). There are also meetings designed specifically to encourage the involvement of young people. In addition, to translate the passive indicators to policy actions, the Boston Indicators Project mobilized hundreds of stakeholders and experts to develop a civic agenda for the Boston area. The agenda has four overarching goals: an open and effective civic culture, world-class human capital, 21st century infrastructure, and 21st century jobs and economic strategies. For each topic area, the agenda includes relevant community indicators, measurable targets, and examples of the actions from various sectors that are contributing to moving the needle (General Accountability Office 2011).

Since 2004, the field of community indicators has been advanced by the Community Indicators Consortium, which serves as a learning network for organizations and individuals interested or engaged in the field of indicators development and application.[4] The consortium provides a forum for those involved in projects covering different topics and geographic scales to share analytic, presentation, and engagement techniques. Convenings make clear that each indicators project strives to move from collecting data for passive display on websites to motivating policy and program improvements. The peer exchange through networks like NNIP and the Community Indicators Consortium offers an important mechanism for showcasing best practices and problem solving around this challenge.

## Data-Driven Action Coalitions

The indicators projects provide periodic broad reviews of community conditions. In contrast, local organizations also use address- and neighborhood-level data in a focused way to guide day-to-day actions of coalitions around specific issue areas. These coalitions may develop communication campaigns for public education, plan coordinated actions, and advocate for policies and investments in their areas.

For example, in the late 2000s, the John D. and Catherine T. MacArthur Foundation supported several property-level information systems to support the preservation of affordable housing. Grantees included the Florida Shimberg Center for Housing Studies, the Ohio Preservation Compact, and the Furman Center for Real Estate and Urban Policy at New York University.[5] The Furman Center, an NNIP partner, coined their system the Subsidized Housing Information Project. The database integrates more than 50 disparate datasets capturing housing units in privately owned buildings receiving federal, state, or local subsidies. The analysts invested time into cleaning the data, combining address-level and parcel-level data and matching properties across subsidy programs (Reina and Williams 2012; US Department of Housing and Urban Development 2013). The data are used to educate the public, support decisionmaking, and inform public debate. To provide broad awareness, the *State of New York City's Subsidized Housing* report presented analysis through maps and summaries by borough and citywide (Begley et al. 2011). A forthcoming analysis will further describe the neighborhood characteristics of subsidized properties. As one example of supporting decisionmaking, the Furman Center analysts annually prepare a list of properties whose subsidies are expiring in the coming years. The City Council then convenes a group of advocates to discuss prioritizing activities for these properties at risk of being removed from the subsidized stock.[6] The data were also used to inform the debate in the 2013 election of the mayor. A series of reports entitled *#NYChousing* drew from the Subsidized Housing Information Project database to present current information and discuss the trade-offs of various policies around affordable housing (Furman Center for Real Estate and Urban Policy 2013). Similar information-focused coalitions emerged in response to the foreclosure crisis in the late 2000s. For example, a brief analyzing trends and spatial patterns of foreclosures in the Washington, DC, area spurred the creation of the Capital Area Foreclosure Network (Pettit et al. 2009). NNIP partner NeighborhoodInfo DC provides Capital Area Foreclosure Network member organizations with detailed information on new properties entering foreclosure. With these weekly data, housing counselors are able to target their efforts precisely, enabling them to reach hundreds of troubled homeowners early to warn them against foreclosure rescue scams and help them determine the best options for their families (Tatian, Grace, and Comey 2013).[7]

While preservation networks focus narrowly on housing, other action coalitions encourage coordination across issue areas. The federal Sustainable Communities Regional Planning grant program spurred locally led cross-sector coalitions to plan and implement strategies supporting regional sustainable and equitable growth. The program, a collaboration among the US Department of Housing and Urban Development, the US Department of Transportation, and the Environmental Protection Agency, supported the grantee networks in identifying how best to target housing, economic, and workforce development and infrastructure investments to spur regional economic activity.[8] From 2010 to 2014, the Department of Housing and Urban Development awarded over $165 million to 74 regional grantees. The program defined six livability principles: (1) provide more transportation choices; (2) promote equitable, affordable housing; (3) enhance economic competitiveness; (4) support existing communities; (5) coordinate policies and leverage investment; and (6) value communities and neighborhoods (US Department of Housing and Urban Development 2012).

The Sustainable Communities Regional Planning program chose proposals that demonstrated the capacity to use data to set and monitor progress toward performance goals and engage stakeholders and residents in meaningful decisionmaking roles. The program specified a common set of performance measures called flagship sustainable indicators. These included overall indicators describing regional populations, but also two measures of equitable development summarizing residents' access to supermarkets and open space within a given distance (US Department of Housing and Urban Development et al. 2012). The program also provided technical assistance to increase the capacity of grantees to use mapping and analysis to guide and track the implementation of their plans. Although the grant program ended in 2014, its fostering of diverse regional partnerships and inclusion of equity as a key component of regional economic progress will leave a legacy in the grantee communities.

Action coalitions focus on many issues beyond housing and transportation. A final example is the Campaign for Grade Level Reading, a collaborative effort by foundations, nonprofit partners, states, and communities "to ensure that more children in low-income families succeed in school and graduate prepared for college, a career, and active citizenship."[9] The campaign was launched with the release of a report sponsored by the Annie E. Casey Foundation that explained that grade-level reading by the end of third grade is a key predictor of academic success (Fiester and Smith 2010). A 2013 research update reviewed evidence showing how living in a high-poverty neighborhood negatively affects school performance. In the results of one study, a child who lives in a high-poverty neighborhood was less likely to graduate from high school, even if he or she was on track on reading tests in third grade (Fiester 2013). Although focused on education outcomes, the coalition recognizes the need for a comprehensive understanding of the child's situation. The Campaign for Grade Level Reading website includes background on how children's health affects school performance and a “starter kit” to help local groups assess health issues and implement health-promotion programs.

More than 100 cities, counties, and regions have signed up to join the Grade Level Reading Communities Network, launching local cross-sector coalitions to craft holistic solutions to improve reading proficiency. To join the network, the community must develop an approved Community Solutions Action Plan, which emphasizes the need for reliable data disaggregated by population group and school. The Community Solutions Action Plan requires the applicant to describe the current situation and trends related to grade-level reading, as well as the larger demographic and economic context. The applicant also commits to having ongoing availability of and access to the data needed to “set baselines, track progress and ensure accountability.”

School absenteeism is one focus area for the campaign. Chang’s essay at the end of this chapter describes four elements of sustainable community–school coalitions for addressing absenteeism: actionable data, capacity building, shared accountability, and positive messaging. She explains how adopting an improved measure—the share of children missing 10 percent or more of school days—enables coalitions to identify the children at most risk better than the use of alternative measures that use absolute days or average days attended. Chang concludes her essay by sharing the implications of chronic absence for neighborhood initiatives.

## Examples of Strategic Use of Neighborhood Data: Case Studies

There are innumerable examples of how neighborhood data are being used to address critical challenges facing metropolitan areas across the nation.[10] In this section, we present three cases that illustrate common themes and innovative approaches that are being used by neighborhood data intermediaries to address concerns in neighborhoods across cities or regions.

### Case Study: Improving Children’s Health in Travis County, Texas

The work of the Austin community in measuring and confronting child obesity in its region demonstrates many of the functions of community information listed in chapter 2, beginning with a presentation of the extent and spatial patterns of an urgent community problem. Their success is due to a combination of sound data shared through clear graphics to inspire community stakeholders and partners to enact a set of targeted interventions to support better children's health outcomes.

The nationwide epidemic of childhood obesity has been well-publicized, and Texas has been one of the hardest-hit states. As of 2007, 20.4 percent of Texas children aged 10 to 17 were obese and 32.2 percent were overweight.[11] This issue disproportionately affects families of low-income and minority populations (Caprio et al. 2008; Singh, Kogan, and van Dyck 2010). Kohl and Cook (2013) review the growing evidence connecting physical fitness, cognitive development, and academic performance. The Austin Independent School District (AISD) conducted its own study of this relationship (Herrera 2009). Across grades 3 to 12, students who were rated as having cardiovascular fitness in the healthy zone had significantly better attendance and higher reading, English language arts, and Texas Assessment of Knowledge and Skills math scale scores than did students with cardiovascular ratings not in the healthy zone.

The first component needed for data-driven action is the creation and availability of the data. The Texas State Senate recognized the importance of understanding children's health status and passed a bill in June 2007 mandating increased physical activity for youth. As part of the implementation of the law, the state required that each district assess the physical fitness of students in grades 3 through 12 each year, including measuring body mass index (BMI) and cardiovascular health[12] (Sage et al. 2010). Physical education instructors collect the data each October and May by administering fitness tests and measuring each child's BMI during class. The decision to create a centralized dataset with this information opened up an exciting opportunity to examine the problem of obesity at a fine-grained level.[13]

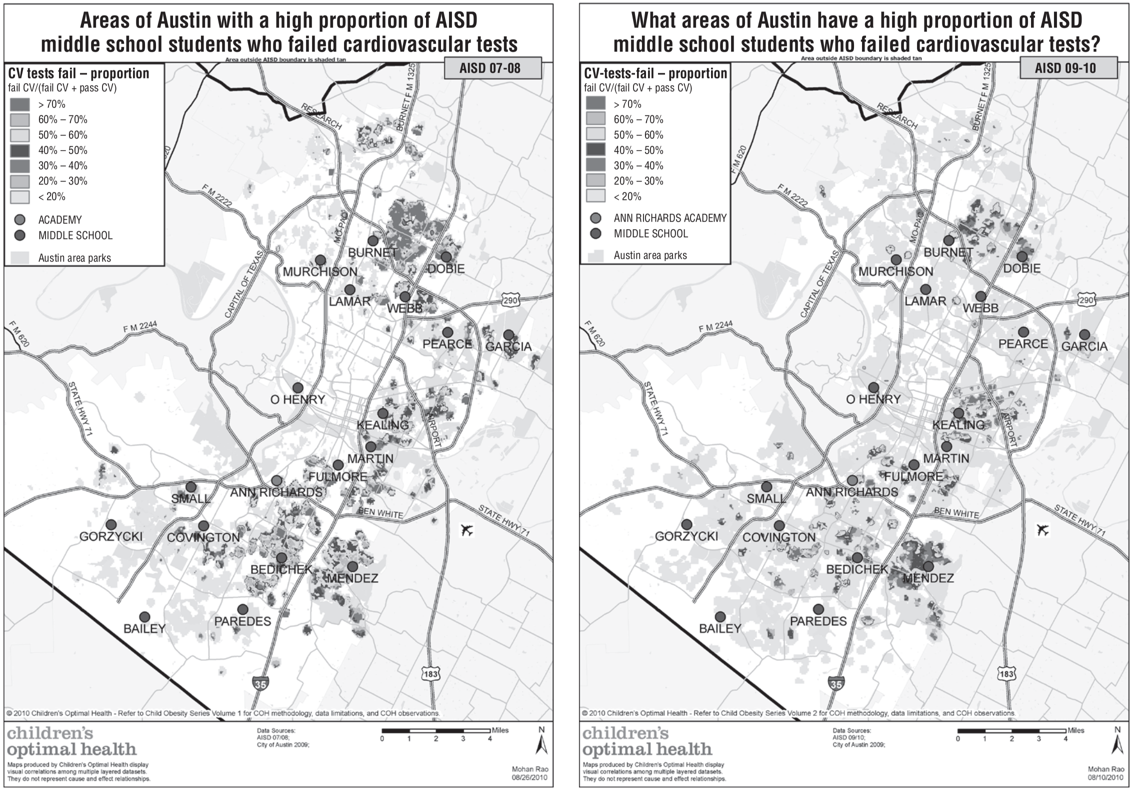
Children's Optimal Health (COH) spearheaded the second stage—moving from data collection to analysis and presentation of the data—to facilitate a common understanding of the problem. COH began as an informal coalition in 2006 to improve children's health in Central Texas and became a registered nonprofit in 2008. The organization's mission is to inform actions that ensure every child in Central Texas becomes a healthy, productive adult. Their board of directors has representatives across sectors (funders; health, education, and social service providers; data providers; nonprofit organizations; for-profit firms; and government agencies) to further their approach of breaking down policy and program silos. Together with their community partners, COH uses shared data, ongoing communication, and collective leadership to further their goals. The organization joined NNIP in 2012.

Beginning in 2008, COH launched a project with AISD to analyze the spatial patterns of overweight and obese children in Austin. Because the organization had been working closely with AISD on other projects, COH was trusted by school officials and had the procedures and agreements in place to handle sensitive data responsibly. COH first produced density maps representing the residences of almost 3,800 middle school students who were overweight, obese, or severely obese. This map had counts of children with weight problems in order to identify the neighborhoods with the greatest number of children in need (see figure 6.1). They also mapped the share of students in an area who were overweight or obese to discern which neighborhoods were disproportionately affected. They found obesity rates for middle schools ranged from 8.6 to 32.1 percent (Sage et al. 2010). COH did similar maps to illustrate the patterns of middle school students with poor cardiovascular health.

COH and its partners recognized that addressing child obesity requires a holistic perspective because it is caused by factors related to individual and family behavior, school curriculum, and neighborhood conditions. The external conditions include a built environment that may encourage or hinder exercise, crime levels that may cause parents to keep their children indoors, and access to healthy foods (Children's Optimal Health 2012). To better understand these contributing factors, COH mapped the health indicators alongside locations of grocery stores, fast-food outlets, and open space. Even among areas with poor health outcomes, the mix of these positive and negative influences varied widely and demonstrated the need for solutions tailored to the neighborhood context.

Next, COH and its partners shared their findings with the wider community to raise awareness about the issue and spur action. In fall 2009, COH convened a Community Summit on Childhood Obesity, which attracted about 100 people, including parents and staff from more than 50 agencies, to review the patterns revealed by the maps and discuss potential policy and programs to intervene. Participants suggested interventions relating to increased access to healthy food and expansion of opportunities for physical activity at home and at school (Sage et al. 2010). Subsequent to this initiative have been efforts by the city and county to improve

**Figure 6.1.** Cardiovascular Rates in Austin Independent School District, 2007–08 (*left*) and 2009–10 (*right*)




*Source:* Analysis of Austin Independent School District data by Children's Optimal Health, 8/26/2010.
*Note:* Color versions of the maps are available at http://www.cohtx.org/download-a-map/

park facilities, improve access to sidewalks and hike and bike trails, and increase access to fresh produce, with special emphasis on neighborhoods of high need as denoted in the maps. Elected officials have used the maps in local policy discussions and referenced them in national presentations. The maps produced by COH were used again when the Michael and Susan Dell Foundation initiated a multiyear place-based initiative in two high-need neighborhoods, extensively engaging community members in decisionmaking to focus on improving physical activity and nutrition.

Over the next several years, COH continued its work mapping the student health indicators. Annual time series maps documented some improved health status in youth. In one example, an area in North Austin experienced unusual improvement in health outcomes from the 2007–08 school year to the 2009–10 school year. At the beginning of the period, more than 70 percent of the middle school children in the area had poor cardiovascular health. By 2009–10, the rates had fallen to less than 50 percent for much of the area. COH learned from the school district staff that the district had piloted HOPSports, a program designed to increase physical activity at the three middle schools that served the neighborhood. COH presented this story and updated analysis for the entire district at a second AISD community summit in November 2010 (Seton Healthcare Family 2010).

Meanwhile, AISD continued to design and implement programs to improve children's health. The AISD School Board included student health as a core value in its strategic plan for 2010 to 2015 (Austin Independent School District 2011). One action step was to "establish goals at each school to prepare children to be healthy, fit, and ready to learn." For example, the plan lists a goal to move the percentage of middle school students with a healthy BMI from 57 percent in 2008–09 to 78 percent by 2011–12 and 100 percent by 2014–15. The plan also includes staged strategies and objectives to meet the stated goals, such as enhanced physical education curriculum and health education in schools (p. 21).

The private sector also participated in the COH coalition and developed new community resources for addressing child health. The Texas Center for the Prevention and Treatment of Childhood Obesity at Dell Children's Medical Center of Central Texas opened a clinic in April 2010 (Dell Children's Medical Center of Central Texas 2010). The clinic built on the success of the Healthy Living, Happy Living/Vida Sana, Vida Feliz multidisciplinary, family-based childhood obesity intervention.

In addition, the philanthropic community played an important role. In late 2009, the RGK Foundation announced a new focus on the "improvement of middle school children's health through physical activity." The foundation commissioned a 2010 report, *Stuck in the Middle: The False Choice between Health and Education in Texas Middle Schools,* which recommended a set of programs and policies (Kelder 2010). In response to a request for proposal issued in September 2012, RGK awarded a $150,000 grant to AISD to enable the implementation of HOPSports and other physical education activities at six middle schools. These additional resources resulted in HOPSports being implemented in all AISD middle schools (Hays Free Press 2012a).

COH's work with AISD attracted the notice of other school districts in Travis County, including Hays Consolidated Independent School District.[14] The steps were similar to those taken in Austin: negotiating a data use agreement, cleaning and analyzing the new data, and sharing the findings with a wide range of stakeholders at a community summit in July 2012. The obesity mapping project was mentioned as one factor in the American School Health Association's selection of Hays Consolidated Independent School District's Superintendent Dr. Jeremy Lyon as the recipient of its 2012 School Health Leadership Award. The award honors superintendents who have fostered school environments that address the connections between health and learning. Like AISD leaders, Dr. Lyon showed leadership in a range of activities to act on the findings of the report, including introducing innovative new programming, expanding physical education activities, and supported the expansion of the school-based student health center (Hays Free Press 2012b).

The need demonstrated by the COH analyses helped to raise external funds to support health programs. The Hemphill Elementary School in the Hays School District, which serves primarily low-income children, was 1 of 12 recipients of a three-year $15,000 "Excellence in Education" Healthy Campus grant, funded by a Texas grocery firm. The grant supported initiatives such as implementing nutrition education throughout the curriculum, including healthier foods in the school meal program and vending machines, creating a community garden, and enhancing physical education classes (Hilsenbeck 2013).

Children's Optimal Health continues to serve the community by analyzing issues at the neighborhood level and sharing the latest news on the factors influencing children's health and potential ways to move the related indicators in the right direction. Achieving improvements in healthy weight and fitness for children is a long-term undertaking, but with the right data, collaborative institutions, and strong leadership in place, the Austin region has begun to see some gains in the indicators. Over a five-year period, the percentage of students in the healthy fitness zone for BMI has increased from 60 to 61 percent; that increase was accompanied by an increase in the percentage of low-income children of about 4 percent. The percentage of students in the cardiovascular healthy zone increased from 62 to 71 percent, but a change in the calculation of aerobic capacity influenced those results. COH analyses and presentation of school and neighborhood indicators served as both a catalyst to action, a tool for targeting interventions, and a way to measure the effectiveness of school and community activities. The relationships that stakeholders have cultivated are now being leveraged to launch further additional projects to help vulnerable children, including building an integrated data system to support case management in programs to reduce family residential and school instability.

## Case Study: Motivating Community Action to Address Neighborhood Disparities in Dallas, Texas

While the Austin actors mobilized their community starting from the single-issue area of children's health, other advocates have scanned multiple aspects of community well-being through a place-based lens. Such was the case in Dallas, where neighborhood indicators presented in a compelling way motivated a coalition of stakeholders, including the *Dallas Morning News,* to work on addressing the disparities between the affluent north side of Dallas and the distressed neighborhoods on the south side.[15]

The Foundation for Community Empowerment, the original NNIP partner in Dallas, was founded by real estate developer Don Williams to facilitate large-scale system change through empowering neighborhood residents. In 2005, the foundation established the J. McDonald Williams Institute as a source of objective research and policy recommendations relevant to urban revitalization and quality of life. The institute's staff noted how the national media helped to elevate the issues of disparity between rich and poor neighborhoods in the wake of Hurricane Katrina in August 2005. Dallas faced similar challenges of limited opportunity in poor neighborhoods, but in the absence of a televised natural disaster, the city's disparities were not getting public attention.

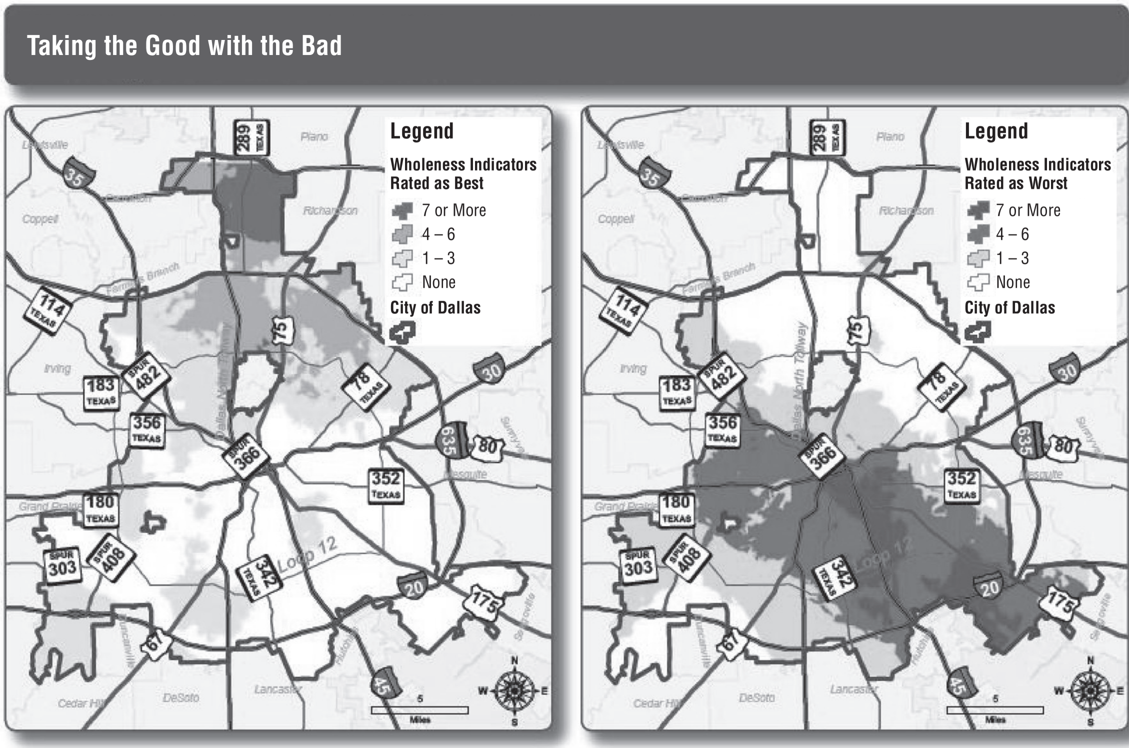
Staff at the J. McDonald Williams Institute believed measurement was the way to raise awareness among a broader audience. By creating a new measure, the Wholeness Index, which captured disparities among different parts of the city, the institute could help define the scope of the problem and track progress as responses were developed (Martin and Bray 2006). The institute defined wholeness as a situation in which "each person in a city enjoys an equally productive and satisfying life, regardless of where in the city he or she lives. In a whole city, residents of every part of town have an equal opportunity to achieve financial success, are equally self-sufficient, and are equally active in political and civic life." The Wholeness Index was composed of 12 individual quality-of-life indicators, including health, housing, education, and other indicators of resident opportunity. Each measure was presented in color-coded maps with explanations of its significance. The disparities in quality of life between the northern parts of the city and the troubled south were consistently pronounced, whether the measures related to jobs, poverty rates, homeownership rates, housing quality, or education. The summary index measure made it easy to communicate the cumulative effects of disadvantage across the various domains (figure 6.2).

The report was updated in 2007 and 2008 to track changes in the indicators (Bray 2007; Williams 2011). For example, the 2008 report showed improvements in the index that reflected better outcomes for South Dallas. Each report was widely disseminated, and the results were highlighted at an annual conference with hundreds of policymakers, researchers, and advocates. In addition to high-profile keynote speakers, panels of experts and practitioners examined specific issue areas.

The report findings and coverage kept the focus on neighborhoods that were historically shut out of opportunity and reinforced the concerns of major community stakeholders like the *Dallas Morning News* (DMN) editorial board, which played a major role in helping to elevate these issues in the city's 2007 mayoral election. With input from qualitative interviews with community members and civic leaders, they formulated candidate questionnaires that focused on issues of disparity and published the results for voters to review. In the fall of that year, the editorial board formally launched the Bridging Dallas North–South Gap project, "a crusade to address the longstanding economic and quality-of-life disparities between the northern and southern halves of the city."[16]

Tom Leppert won the mayoral election. His address at the 2008 Williams Institute conference recognized the importance of understanding

**Figure 6.2.** Wholeness Index in Dallas, 2006




*Source:* Martin and Bray (2006).
*Note:* Color versions of the maps are available at http://www.neighborhoodindicators.org/library/catalog/dallas-whole-city-2006-wholeness-index-summary-report

patterns of disadvantage in the city and the urgency of addressing the gaps. As one strategy to accomplish this, the mayor instituted the Mayor's Southern Dallas Task Force,[17] which divided Southern Dallas into 10 neighborhoods and brought citizens and business leaders together over two years to suggest plans for improvement in each neighborhood. Strategies ranged from smaller steps, like neighborhood branding campaigns and new security cameras in business areas, to larger ones, like reopening a hospital or bringing more retail to the neighborhoods.

Despite raised expectations, the task force recommendations got little attention from policymakers (University of Texas at Dallas 2008; *Dallas Morning News* 2009). The DMN staff planned to call on the city government to seriously consider the recommendations of the task force, but instead of a stand-alone editorial, they decided to reinforce their argument with a rich portrait of South Dallas. This special section would include neighborhood conditions and advocate for policies and programs to improve conditions for residents there. The DMN staff approached the authors of the Wholeness Index to assemble and analyze many sources of data to illustrate the disparities between northern and southern Dallas. (The Williams Institute had recently been renamed the Institute of Urban Policy and Research and moved to the University of Texas at Dallas.) The researchers worked closely with DMN staff to identify five focus neighborhoods in South Dallas and provided information for each neighborhood that described their demographics, health conditions, crime, and education outcomes.

The Institute of Urban Policy and Research team saw the limitations of the secondary data and also conducted a two-month windshield survey. The institute trained members of its Community Research Team, who were residents of South Dallas communities, to systematically collect information about land use, housing conditions, walkability, and other quality-of-life factors. These data were then assembled into a mapping program that allowed the exploration and summary of these physical characteristics by neighborhood.

The institute's work with DMN resulted in a major eight-page spread in a Sunday paper in September 2009, complete with maps, data, and stories about the five South Dallas neighborhoods. The collection of articles documented current disparities, highlighted programs that were helping to close the gap, and called for additional investments in the neighborhoods. The impressive feature section earned the DMN editorial staff a Pulitzer Prize in 2010 for editorial writing (Wilkins 2010).

DMN is committed to keeping the city government accountable for progress in reducing disparities. They continue to use data and information to drive work in the editorial and news rooms. For example, the editorial staff has a monthly feature, *10 Drops in the Bucket,* which spotlights specific persistent problems in South Dallas (such as a dangerous vacant property with code violations or traffic safety issues). The writers call attention to these problems in a very public way and monitor them for signs of progress.

Urged on by the continued DMN coverage and ongoing research from the Institute of Urban Policy and Research, elected leaders carry on efforts to revitalize South Dallas. Mayor Mike Rawlings, elected in 2011, launched Grow South in April 2012. This city-led initiative aims to bring attention to the many assets and civic pride in Southern Dallas communities. The 10 goals for the initiative range from engaging residents in these neighborhoods to strengthening schools, educating the public about assets in South Dallas, and investing in economic development. The Grow South program is building public–private partnerships to bring resources and economic development to the city's southern neighborhoods.[18] Although still early in the initiative, the city appears to be leveraging neighborhood indicators to hold themselves accountable for making progress. The program's first-year report included profiles of eight focus areas, reporting on measures of commitment, resilience, and amenities for 2011 and 2012 (City of Dallas 2013).

In May 2013, the City of Dallas and Grow South joined with Dallas Area Habitat for Humanity, Safer Dallas Better Dallas, and Communities Foundation of Texas to create EPIC (Economic Partners Investing in Communities) Dallas (Bush 2013).[19] The organizations share common goals of increasing public safety, access to homeownership opportunities, and economic development. EPIC Dallas has created a board that will study the work of the member organizations and identify areas of overlap that could be strengthened by better coordination and joint fundraising. EPIC Dallas has identified a set of neighborhood indicators to track the initiative's progress, including property values, violent and property crime rates, employment rates, educational attainment, voter registration, homeownership rates, and the number of neighborhood associations and crime watch groups. Its first activities have raised $650,000 to support targeting crime hotspots in five neighborhoods (Hallman 2013).

The Dallas community faces continuing challenges in overcoming decades of disinvestment in its southern neighborhoods. With the release of the first Wholeness Index, neighborhood indicators were used to highlight the extent of the inequities in city and broaden support for tackling the problem. Progressive DMN leaders helped keep the issue in the public spotlight and tied it to specific calls for action. Philanthropic groups and government agencies followed suit, increasingly using data to target their work and measure their progress.

The Dallas case study shows the power of having a data and intermediary organization like the Institute of Urban Policy and Research to support forward-thinking government and community leaders. The Dallas region's journey from the mid 2000s to today illustrates how all sectors can learn to use small-area data to advance the continuing agenda for neighborhood improvement and social change.

## Case Study: Improving Health Care Systems in Camden, New Jersey

Our final case echoes some themes of the first two cases but also shows how analysis of detailed data can support interventions at the neighborhood and individual levels. Because of the strong legal protections of HIPAA and because health system data are proprietary, few local data intermediaries have small-area health data. Thanks to an entrepreneurial doctor and progressive hospital leadership, the Camden community has surmounted these difficulties and shown how local data can motivate collaboration and inform action to improve the health of its residents.

For a relatively small place, Camden has an unusual set of health resources: Three hospitals inside the city serve as anchor institutions and strengthen civic leadership. For example, Virtua Health System was a supporter in the 2001 creation of the NNIP partner organization CamConnect. The year following the founding of CamConnect, primary care providers began meeting informally to discuss their experiences in serving Camden residents. Led by Dr. Jeffrey Brenner, they decided to formalize their relationship by creating a nonprofit organization called the Camden Coalition of Health Care Providers. Its mission was to build an integrated health delivery model to provide better care for Camden residents.[20]

Dr. Brenner, the organization's executive director, understood the potential for data to inform community problem solving and had already been looking at hospitalization and crime data to identify patterns of violence in the city (Gawande 2011). All three of the city's hospitals (Cooper, Our Lady of Lourdes, and Virtua) also wanted to understand their patients better and were concerned about the rising costs of health care for uninsured people. They agreed to share their patient billing data, including emergency department visits and inpatient hospitalizations, with CamConnect and the Camden Coalition of Health Care Providers. Negotiations to finalize the data transfer took almost a year, and the organizations put extensive protections in place to comply with HIPAA and the health systems' institutional review boards (CamConnect 2010). The data contained the address, diagnosis codes, demographic information, and financial information about the hospital visit. The analysts faced many challenges in merging data across the three hospital data systems, including developing probabilistic algorithms to match individuals who visited multiple hospitals. Their first groundbreaking analysis of 2003 data was published in 2006 (CamConnect and Camden Coalition of Health Care Providers 2006). They discovered that half the population in the city used hospital services in 2003. Their analysis helped them understand basic descriptive information, such as the incidence and rates of chronic disease and preventable incidents, the frequency of visits, and the distribution of payment methods. Their maps of the rates of hospital visits due to diabetes, falls, drug abuse, and assaults demonstrated the variation across the city. The coalition recognized that all four of these causes of hospital visits could be addressed by community intervention.

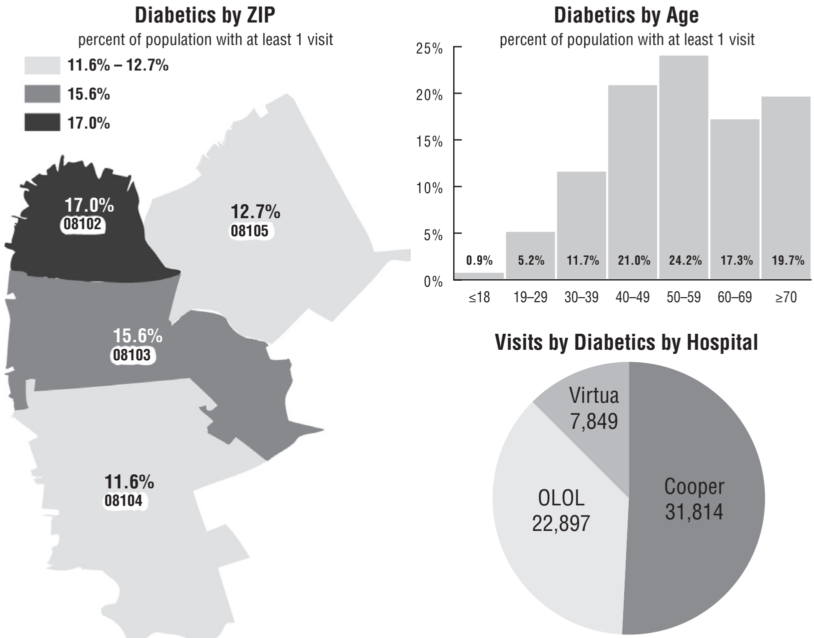
The next stage of the analysis expanded the coverage to six years. The coalition identified more than 7,000 individuals with type 2 diabetes who visited Camden emergency departments or hospitals 62,000 times between 2002 and 2008 and accumulated charges of over $1.5 billion (see figure 6.3). Just knowing the facts was not enough to improve the situation for these patients. With support from the Merck Company's Foundation, the coalition launched the Citywide Diabetes Collaborative to advance proactive, comprehensive care for people suffering from the disease.[21] The maps showing the areas where diabetes was most prevalent helped the collaborative to focus their outreach efforts for classes on how to better manage the illness and reduce the need for hospital care.

The analysis also demonstrated that a small share of the patients generated most of the costs and visits. From 2002 to 2007, 20 percent of the patients accounted for 90 percent of the costs.[22] In 2007, Brenner worked with social workers and emergency room doctors to identify the most frequent visitors, termed the *superutilizers.* Health practitioners visited these patients at home to learn more about how social and environmental factors contributed to the health conditions and the patient's ability to follow up with medical recommendations. Patients were also given a

**Figure 6.3.** Diabetes in Camden, New Jersey, 2002–2008

## Diabetes in Camden

Data from a Camden City Comprehensive Health Database (July 2002 – June 2008)




| Number of Visits | Total Visits | Total Patients | Charges | Receipts |
|---|---|---|---|---|
| 1 to 10 | 22,075 | 5,270 | $674,344,336.42 | $92,597,330.57 |
| 11 to 20 | 16,134 | 1,126 | $431,920,873.45 | $55,148,779.38 |
| Over 20 | 24,351 | 645 | $444,163,826.50 | $55,970,659.88 |
| TOTAL | 62,560 | 7,041 | $1,550,429,036.37 | $203,716,769.83 |

Camden Coalition of Healthcare Providers
www.camdenhealth.org

Version 3
6-12-09

Linking Communities with Information
CAMConnect

*Source:* Camden Coalition of Healthcare Providers and CamConnect.

dedicated phone number to call for medical advice. The coalition documented the improvements for the first 36 participants, with a 40 percent reduction in hospital visits. By late 2010, Dr. Brenner's team had provided care for more than 300 people (Gawande 2011).

The coalition has hired in-house analysts to update and manage the data and has routinized much of its analysis work since the first analysis in 2006. The Health Information Exchange, launched in 2010 with the original three health systems, allows hospital data to be viewed by over 100 healthcare providers in Camden. The exchange's real-time data feeds from health care facilities inform clinicians' treatment decisions. The exchange also supports the coalition's continuing work to identify the patients who could most benefit from their intervention program. Each morning, the coalition staff generates a list identifying Camden residents who were admitted to a city hospital the previous day and have histories of hospitalizations for chronic conditions.[23]

Through the coalition's journey over the past decade, the members are able to share the lessons they have learned about strategic coordination, the importance of communication, and the need for sustainability. They identified the motivations that brought players to the table: a mutual dissatisfaction with costs and poor outcomes from status quo practices in the health care system.

Two of the lessons in particular echo the experiences of NNIP and the previous case studies (Camden Coalition of Healthcare Providers 2014):

- No single organization or health system has enough resources to solve a citywide problem. Solutions require collaboration among all stakeholders to provide adequate support for implementation.
- Initiatives need to be data-driven in order to direct the most effective strategies and to understand where to dedicate resources.

In 2013, the MacArthur Foundation recognized Dr. Brenner's innovative ideas by awarding him their Genius grant. The coalition is moving in several directions to advance their practice and the field as a whole. After building trust as responsible users of the hospital claims data, they have begun to expand the data warehouse by incorporating additional sources, including lab results, neighborhood-collected BMI data, and birth records, and by expanding coverage to suburban hospitals.[24]

They are also helping other communities that are trying to develop similar systems through technical assistance, including developing a hot-spotting toolkit with the support of the Commonwealth Foundation.[25] The coalition also received a 2014 John S. and James L. Knight Foundation News Challenge grant to develop an open-source health care dashboard in partnership with a technology development firm. This system will enable other communities to organize and upload their own data (O'Connor 2014). The system will aggregate the individual-level data by demographic group and geography so that community groups and policymakers alike can better understand patterns of health care use through a publicly accessible online site. Even with technical assistance on data development, community process, and technology, challenges to expanding the practice remain, including hospital bureaucracy, federal financing practices, and entrenched institutional interests (Blumgart 2012). Nonetheless, opportunities for smarter use of data in health care will increase as the Affordable Care Act spurs more and more health records to be moved to electronic and integrated systems. The Camden Coalition of Health Care Providers offers an inspirational model of how these data can both improve care for individual patients and be used by stakeholders to develop community-level interventions.

## Conclusion

The examples and case studies from this chapter reveal the ingredients needed to overcome the challenges in getting data used in communities. Each organization was locally based and participating in civic conversations about issues facing the community. Each took advantage of newly available data or presented data in a fresh way. The organizations carefully cleaned and tabulated the data, but they recognized that statistics and colorful maps were insufficient. The capacity and persistence of local institutions and individual champions were essential to any progress. The information needed to be shared with the groups, whether inside or outside of government, that had the means and motivation to address the problem. Most of the examples cover multiple years, showing that the process of assembling a shared understanding of a problem and the appropriate response is slow. Direct cause-and-effect relationships between the data presentation and community change are difficult to trace, but the examples demonstrate how neighborhood indicators provided a catalyst for community concern and action.

The case studies also show how the functions of community indicators introduced in chapter 4 work in combination. In every instance, the work began by careful analysis of data to capture the scope and distribution of problems and resources (situation analysis). The data were then used to engage and mobilize stakeholders to seek solutions (education and advocacy). Coalitions were built to plan and implement programs as well as develop policy (policy analysis and planning). And in several of the cases they began to use data to evaluate the success of the efforts and make improvements as needed (performance management and evaluation). In real experience, these functions are seldom distinct; they tend to emerge in tandem and repeat themselves through deeper use of the data as successful projects move to maturity. The field has now developed sufficiently that various cases such as those described above have come full circle and can illustrate the powerful role of neighborhood data.

## NOTES

1. The power of data to inspire collaborative problem-solving across sectors is one of the principles of the NNIP cross-site projects. See Kingsley and Hendey (2010) and LaVigne, Cowan, and Brazzell (2006) for examples related to school readiness and prisoner reentry, respectively.

2. For more information, see the Kirwan Institute website at http://kirwaninstitute.osu.edu.

3. Meg Merrick, "Defining, Tracking, and Displaying Regional Equity Conditions: Two Approaches from the Portland-Vancouver Metropolitan Region," a paper prepared for the Urban Affairs Association Conference, March 2013.

4. For more information, see the Community Indicators Consortium website at http://www.communityindicators.net.

5. The chapter 2 essay by Wascalus and Matson describes HousingLink's Streams, a similar subsidized housing catalog for Minnesota. The NNIP Partner Community Research Partners also support the Ohio Preservation Compact (Foreclosure-Response.org. 2009). Finally, Kingsley and Pettit (2008) describe the Coalition for Affordable Housing in the District of Columbia, a locally funded program parallel to the Subsidized Housing Information Project in New York City.

6. Mary Weselcouch, e-mail communication with Kathryn L. S. Pettit, August 28, 2013.

7. Similar programs existed in Kent County, Michigan, supported by the Community Research Institute at Grand Valley State University (Chartkoff and Rotondaro 2008) and in Chicago through the Regional Home Ownership Preservation Initiative (http://www.regionalhopi.org/; accessed January 12, 2014) supported by the Woodstock Institute.

8. Three NNIP partners (the Chicago Metropolitan Agency for Planning, the Mid-America Regional Council, and the Metropolitan Area Planning Council in Boston) were Sustainable Communities Initiative grantees. Several other partners (Minneapolis, Minnesota; Denver, Colorado; and St. Louis, Missouri) provided analytic support for their local grantee. For more examples of the role of neighborhood data in the initiative, see the NNIP website at http://neighborhoodindicators.org/issue-area/209.

9. Campaign for Grade Level Reading website. Accessed January 20, 2014. http://gradelevelreading.net/.

10. See Cowan (2007) and Kingsley and Pettit (2011) for other case studies on data use by the organizations in NNIP.

11. See Texas Comptroller of Public Accounts, "Obesity and Overweight Prevalence among Children Aged 10 to 17, by State, 2003 vs. 2007," at http://www.window.state.tx.us/specialrpt/obesitycost/exhibits/ex4.php. Accessed January 12, 2014.

12. As of 2009, 13 states had legislation requiring school-based BMI measurement programs, with implementation at various stages (Nihiser et al. 2009).

13. For a series of essays exploring the early data from the Texas Fitness Testing from a variety of perspectives, see the September 2010 Special Supplement to *Research Quarterly for Exercise and Sport* on the Findings of Texas-Wide School-Based Fitness Testing.

14. COH completed a similar study for Manor Independent School District northwest of Austin in May 2013.

15. This case study was written with the assistance of Leah Hendey of the Urban Institute and Timothy Bray at the Institute of Urban Policy Research at the University of Texas at Dallas.

16. For additional information, see http://www.dallasnews.com/opinion/north-south-dallas-project/.

17. For additional information, see http://www.southerndallas.org/.

18. For additional information about Grow South, see http://www.dallascityhall.com/government/mayor/growSouth.html.

19. For additional information, see "Dallas Civic and Philanthropic Leaders Announce Groundbreaking Epic Initiative." Press release, May 14, 2013. http://www.dallasareahabitat.org/c/document_library/get_file?p_l_id=33636&folderId=48594&name=DLFE-2215.pdf.

20. For additional information about the history of Camden Coalition of Health Care Providers, see http://www.camdenhealth.org/about/about-the-coalition/history/.

21. For additional information about the Camden diabetes collaborative, see http://www.camdenhealth.org/programs/citywide-diabetes-collaborative/.

22. Derek Ziegler, "Overview of Camden Emergency Department and Hospitalization Data." National Neighborhood Indicators Partnership webinar, August 21, 2008.

23. For additional information about the Camden Health Information Exchange, see http://www.camdenhealth.org/programs/health-information-exchange.

24. Truchil, Aaron. Telephone interview with Kathryn Pettit. January 23, 2014.

25. For additional information about the Hot Spotting Toolkit, see https://www.aamc.org/initiatives/hotspotter/toolkit/.

## REFERENCES


Austin Independent School District. 2011. *Strategic Plan 2010–2015.* Austin, TX.

Begley, Jaclene, Caitlyn Brazill, Vincent Reina, and Max Weselcouch. 2011. *State of New York City's Subsidized Housing.* New York: New York University. September.

Benner, Chris, and Manuel Pastor. 2012. *Just Growth: Prosperity and Inclusion in America's Metropolitan Regions.* Routledge Press.

Blumgart, Jake. 2012. "Urgent Care: Recovering from the Urban Hospital Addiction." *Forefront* 35: December 12. http://nextcity.org/forefront/view/urgent-care.

Bray, Timothy. 2007. *Is Dallas Becoming More Whole?* Dallas, TX: J. Williams McDonald Institute.

Bush, Rudolph. 2013. "Umbrella Group Will Coordinate Efforts to Help Dallas' Needy Neighborhoods." *Dallas Morning News.* May 14.

CamConnect. 2010. *Geography and Epidemiology of Health Data for the City of Camden: Creating a Neighborhood Health Indicator Database.* Camden, NJ: CamConnect.

CamConnect, and Camden Coalition of Health Care Providers. 2006. "Camden Hospital and Emergency Room Health Data 2003." Camden, NJ: CamConnect.

Camden Coalition of Healthcare Providers. 2014. www.camdenhealth.org/about-the-coalition/our-mission/

Caprio, Sonia, Stephen R. Daniels, Adam Drewnowski, Francine R. Kaufman, Lawrence A. Palinkas, Arlan L. Rosenbloom, and Jeffrey B. Schwimmer. 2008. "Influence of Race, Ethnicity and Culture on Childhood Obesity: Implications for Prevention and Treatment: A Consensus Statement of Shaping America's Health and the Obesity Society." *Diabetes Care.* 31 (11): 2211–21.

Chartkoff, Shelby, and Gustavo Rotondaro. 2008. *Sold Short: Residential Foreclosures in Kent County, 2004–2007.* Grand Rapids, MI: Grand Valley State University.

Children's Optimal Health. 2012. *Child Obesity by Neighborhood, 2007–2010.* Austin, TX: Children's Optimal Health.

City of Dallas. 2013. *Grow South Community Indicators: One Year Progress Report.* Dallas, TX: Office of Economic Development, City of Dallas. http://www.dallascityhall.com/government/mayor/pdf/2013_FocusAreasAnnualReport.pdf.

Cowan, Jake. 2007. *Stories: Using Information in Community Building and Local Policy.* Washington, DC: Urban Institute.

*Dallas Morning News.* 2009. "Task Force Deferred." September 20.

Dell Children's Medical Center of Central Texas. 2010. "Texas Center for the Prevention and Treatment of Childhood Obesity Opens Clinic." April 22. http://www.dellchildrens.net/about_us/news/2010/04/22/texas_center_for_the_prevention_and_treatment_of_childhood_obesity_opens_clinic.

Duncan, G. J., and R. Murnane. (2011). *Whither opportunity? Rising inequality, schools, and children.* New York: Russell Sage Foundation.

Ellen, Ingrid Gould, and Margery Austin Turner. 1997. "Does Neighborhood Matter? Assessing Recent Evidence." *Housing Policy Debate* 8: 833–866.

Fernandez, R. M., and C. Su. (2004). Space in the Study of Labor Markets. *Annual Review of Sociology, 30,* 545–569.

Fiester, Leila. 2013. *Early Warning Confirmed: A Research Update on Third-Grade Reading.* Baltimore: Annie E. Casey Foundation.

Fiester, Leila, and Ralph Smith. 2010. *Early Warning: Why Reading by the End of Third Grade Matters: a KIDS COUNT Special Report.* Baltimore: Annie E. Casey Foundation. May.

Foreclosure-Response.org. 2009. *Case Study: Ohio Foreclosure Prevention Task Force.* Accessed January 12, 2014. http://www.foreclosure-response.org/policy_guide/coordinated_response.html?tierid=310.

Friedrichs, J. R., G. Galster, and S. Musterd. (2003). Neighbourhood Effects on Social Opportunities: The European and American Research and Policy Context. *Housing Studies, 18*(6), 797–806.

Furman Center for Real Estate and Urban Policy. 2013. *#NYChousing: 10 Issues for NYC's Next Mayor.* New York: New York University.

Gawande, Atul. 2011. The Hot Spotters. *New Yorker.* January 24.

General Accountability Office. 2011. *Key Indicator Systems: Experiences of Other National and Subnational Systems Offer Insights for the United States.* Washington, DC: General Accountability Office.

Hallman, Tristan. 2013. "Dallas Police Booster Groups Donate $350,000 for Crime-Fighting Technology." Safer Dallas. October 8.

Hays Free Press. 2012a. School Briefs: May 9, 2012. http://www.haysfreepress.com/2012/05/09/school-briefs-may-09-2012/.

Hays Free Press. 2012b. School Briefs: October 10, 2012. http://www.haysfreepress.com/index.php/local-news/hays-county/86-schools/14066-school-briefs-oct-10-2012.

Herrera, Angelica Ware. 2009. *E-Team Report: DPE Publication Number 07.95.* October. Accessed January 12, 2014. http://curriculum.austinisd.org/pe_health/fitnessgram/documents/FINAL_E-TeamCardiovascularTAKSandAttendance.pdf.

Hilsenbeck, Kim. 2013. "Hemphill Wins $15K Healthy Campus Grant." Hays Free Press. February 6. Accessed September 2, 2013. http://haysfreepress.com/index.php/schools?start=80.

Kelder, Stephen. 2010. *Stuck in the Middle: The False Choice Between Health and Education in Texas Middle Schools. Recommendations for Programs and Policies.* Austin, TX: RGK Foundation. October.

Kingsley, G. Thomas, and Leah Hendey. 2010. *Using Data to Promote Collaboration in Local School Readiness Systems.* Washington, DC: Urban Institute.

Kingsley, G. Thomas, and Kathryn L. S. Pettit. 2008. *Data and Decisions: Parcel-Level Information Changing the Way Business Gets Done.* Washington, DC: Urban Institute.

Kingsley, G. Thomas, and Kathryn L. S. Pettit. 2011. "Quality Of Life at a Finer Grain: The National Neighborhood Indicators Partnership." In *Community Quality-of-Life Indicators: Best Cases V.* Springer Publishing.

Kohl, Harold W., III, and Heather D. Cook. 2013. *Educating the Student Body: Taking Physical Activity and Physical Education to School.* Washington, DC: National Academies Press.

LaVigne, Nancy G., Jake Cowan, and Diana Brazzell. 2006. *Mapping Prisoner Reentry: An Action Research Guidebook.* Washington, DC: Urban Institute.

Leventhal, Tama, and Jeanne Brooks-Gunn. 2000. "The Neighborhoods They Live In: The Effects of Neighborhood Residence Upon Child and Adolescent Outcomes." *Psychological Bulletin* 126: 309–337.

Martin, Marcus, and Timothy Bray. 2006. *Is Dallas a Whole City?* Dallas, TX: J. Williams McDonald Institute.

Mueller, E. J., and J. R. Tighe. (2007). Making the Case for Affordable Housing: Connecting Housing with Health and Education Outcomes. *Journal of Planning Literature, 21*(4), 371–385.

Nelson, C. A., and M. Sheridan. (2009). Lessons from Neuroscience Research for Understanding Causal Links between Family and Neighborhood Characteristics and Educational Outcomes. In G. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools and children's life chances* (pp. 27–46). New York: Russell Sage.

Nihiser, Allison J., Sarah M. Lee, Howell Wechsler, Mary McKenna, Erica Odom, Chris Reinold, Diane Thompson, and Larry Grummer-Strawn. 2009. "BMI Measurement in Schools." *Pediatrics* 124 (Supplement 1): S89–S97.

O'Connor, Patrick. 2014. "Obama Data Analysts Study Health-Care Habits under Obama care." *Wall Street Journal.* January 15. http://blogs.wsj.com/washwire/2014/01/15/obama-data-analysts-study-health-care-habits-under-obamacare/.

Pettit, Kathryn L. S., Mary Cunningham, G. Thomas Kingsley, Leah Hendey, Jennifer Comey, Liza Getsinger, and Michel Grosz. 2009. *Housing in the Nation's Capital 2009.* Washington, DC: Fannie Mae.

Reina, Vincent, and Michael Williams. 2012. "The Importance of Using Layered Data to Analyze Housing: The Case of the Subsidized Housing Information Project." *Cityscape: A Journal of Policy Development and Research* 14 (1) March.

Sage, William M., Matthew Balthazar, Steven Kelder, Susan Millea, Stephen Pont, and Mohan Rao. 2010. "Mapping Data Shape Community Responses to Childhood Obesity." *Health Affairs* 29 (March): 498–502.

Sampson, R. J., J. D. Morenoff, and T. Gannon-Rowley. (2002). Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual Review of Sociology, 28*(1), 443–478.

Seton Healthcare Family. 2010. "Finding a Fitter Future: Children's Optimal Health Report Reveals Improved Middle School Fitness." September 16. https://www.seton.net/about_seton/news/2010/09/16/finding_a_fitter_future_childrens_optimal_health_report_reveals_improved_middle_school_fitness_2.

Shinn, M., and S. M. Toohey. (2003). Community Contexts of Human Welfare. *Annual Review of Psychology, 54*(1), 427–459.

Singh, Gopal K., Michael D. Kogan, and Peter C. van Dyck. 2010. "Changes in State-Specific Childhood Obesity and Overweight Prevalence in the United States from 2003 to 2007." *Archives of Pediatrics and Adolescent Medicine* 164 (7): 598–607.

Tatian, Peter A., Rebecca Grace, and Jennifer Comey. 2013. *Making Data Accessible: Foreclosure Tracking in Washington, DC.* Washington, DC: Urban Institute. July.

University of Texas at Dallas. 2008. "Institute Releases Annual Dallas Wholeness Index." October 22. Accessed November 7, 2013. http://www.utdallas.edu/news/2008/10/22-004.php.

US Department of Housing and Urban Development. 2012. *Office of Sustainable Housing and Communities: Helping Communities Realize a More Prosperous Future.* Washington, DC: US Department of Housing and Urban Development.

US Department of Housing and Urban Development. 2013. *How Research Tools Are Assisting Communities to Preserve, Plan Affordable Housing. Evidence Matters.* Summer. http://www.huduser.org/portal/periodicals/em/summer13/highlight2.html#title.

US Department of Housing and Urban Development, Interagency Partnership for Sustainable Communities, ICF International, and Summit Consulting. 2012. *Office of Sustainable Housing and Communities Guidance on Performance Measurement and Flagship Sustainability Indicator Flag Sheets.* Washington, DC: US Department of Housing and Urban Development.

van Ham, M., D. Manley, N. Bailey, L. Simpson, and D. Maclennan. (2012). Neighbourhood effects research: New perspectives. In M. van Ham, D. Manley, N. Bailey, L. Simpson, & D. Maclennan (Eds.), *Neighbourhood Effects Research: New perspectives* (pp. 1–21). New York: Springer.

Weiss, C. C., M. Purciel, M. Bader, J. W. Quinn, G. Lovasi, K. M. Neckerman, et al. (2011). Reconsidering Access: Park Facilities and Neighborhood Disamenities in New York City. *Journal of Urban Health, 88*(2), 297–310.

Wilkins, Chris. 2010. "Dallas Morning News Wins Pulitzer for Editorial Writing." *Dallas Morning News.* April 13. http://www.dallasnews.com/news/community-news/dallas/headlines/20100413-Dallas-Morning-News-wins-Pulitzer-for-5971.ece?nclick_check=1.

Williams, Ashley. 2011. "Dallas Metropolitan Area: Weathering the Storm and Tackling Disparities." *MetroTrends Blog.* Washington, DC: Urban Institute. http://metrotrends.org/spotlight/dallas.cfm.

ESSAY

# Creating and Using Neighborhood Data to Shape a More Equitable Regional Greenspace Policy

Meg Merrick and Sheila Martin

Can neighborhood-level data influence a community's understanding of social equity? Can it leverage that understanding to influence regional policy? This essay demonstrates that neighborhood-level data improve our understanding of social equity issues as they play out across the landscape. Furthermore, neighborhood-level data can be a powerful tool for advocating for policies that address these inequities.

We focus on the development and use of the Regional Equity Atlas, a neighborhood-level spatial analysis of equity issues in the Portland, Oregon, metropolitan region. The Coalition for a Livable Future (CLF), a nonprofit coalition of community-based groups based in Portland, published the Regional Equity Atlas in 2007, with significant leadership from the Portland Audubon Society, a founding member of CLF (Coalition for a Livable Future 2007). This essay explains how CLF and Audubon recognized the potential of these data for advancing their advocacy, used the neighborhood data and maps—even before the publication of the Atlas—to influence regional greenspace policy, and successfully advocated for a bond measure that provided funding and encouraged partnerships to address park-deficient areas.

The Regional Equity Atlas continues to shape the resulting parks funding program by directing outreach and funding to projects that offer opportunities to increase access to parks, greenspaces, and nature for low-income communities and communities of color. Ultimately, the inequities revealed by the Atlas influenced Audubon's outreach efforts and the direction of their long-term strategic plan.

## Demonstrating the Power of Neighborhood Data

Before launching its effort to create the Regional Equity Atlas, CLF experienced the power of using objective, credible data for advocacy through its efforts to influence the regional framework plan. This plan, originally adopted by Metro (Portland's regional government) in 1997, unites all of Metro's adopted land use planning policies and requirements into a single document; thus, it has enormous influence on the future of the region. The framework plan influences investments in regional transportation, the placement or expansion of the urban growth boundary, protection of areas outside the urban growth boundary, housing densities, urban design, parks and open space, and the management of water resources. The framework plan's expansive set of issues reflected the holistic view of community development that spawned the coalition and invited the type of cross-issue advocacy that CLF's broad coalition required. The coalition focused its advocacy for the framework plan on an all-encompassing notion of livability that created the banner under which the varied interests of the member organizations (affordable housing, the environment, urban design, health, and public transportation, to name a few) could act as a united front in the advocacy arena. Recognizing the importance of objectively derived evidence to persuasive argument in the policymaking arena, the coalition bolstered its case by referencing data, analyses, and maps that it commissioned from Myron Orfield of the University of Minnesota and later published, in 1998, as *Portland Metropolitics.*

This approach—uniting disparate groups to advocate for common positions and using maps, data, and analysis to support those positions—proved highly successful. Mary Kyle McCurdy, staff attorney at 1000 Friends of Oregon (also a coalition founding member organization) estimated that of 15 pages of recommendations related to land use, transportation, and housing presented by CLF to the Metro Council, two-thirds were adopted into the framework plan that would guide the future growth of the Portland region (Walljasper 1999).

## The Regional Equity Atlas

In the aftermath of its victories on the regional framework plan, CLF focused on implementing the adopted regional policies and began to frame its work more explicitly in terms of sustainability, defined as was popular at the time, as the three Es: the environment, the economy, and equity. It was during this period that the coalition saw an opportunity to make a unique and substantial contribution to the regional policy discussion by focusing on equity, which was thought to be the least understood of the three Es. Moreover, CLF believed that the kinds of smart growth policies that had been adopted in Metro's regional framework plan (with CLF's encouragement) weren't adequate to addressing the inequities in the region.

In addition, CLF wanted to build on Orfield's *Metropolitics* analysis to create an approach that would harness the analytic and graphic power of mapping to construct an "equity atlas." Thus, in 2003, CLF, in partnership with the Institute of Portland Metropolitan Studies and the Population Research Center at Portland State University, launched a ground-breaking initiative to create the nation's first regional equity atlas. CLF saw the equity atlas as an opportunity to build on its success in using objective information to advocate for livability causes. Portland State University provided credibility as a neutral provider of data and technical capabilities.

### Defining Equity

Equity is an amorphous notion that requires definition. The definition that CLF developed was derived through a participatory process among its members and community allies and leaders (particularly targeting communities of color) that resulted in a trio of aspirations for the region. An equitable region would be one in which (1) all residents would have access to opportunities for meeting their basic needs and advancing their health and well-being (access to good jobs, transportation choices, safe and stable housing, a good education, quality health care, a range of parks and natural areas, vibrant public spaces, and healthful foods); (2) the benefits and burdens of growth and change would be fairly shared among the region's communities; and (3) all the region's residents and communities would be fully involved as equal partners in public decisionmaking (CLF 2007). The Regional Equity Atlas was intended to create a baseline assessment of the region against these aspirations that would be updated over time.

## Defining Access

Although access, which is central to CLF's first aspiration, has implications that go far beyond physical distance, the bulk of the analytical work of CLF's Regional Equity Atlas focused primarily on the geographic proximity of populations (particularly those who are poor and/or belong to communities of color) to various community assets. The Atlas measured the benefits and burdens of growth and change in terms of the percentage or numeric change in the distribution of people and assets over time and across space. The Atlas did not address the third aspiration, which emphasizes participatory democracy, because of the inadequacy of data to produce meaningful results in the spatial realm. Fundamentally, Atlas analyses target the spatial manifestations of the equity issue as defined by CLF and don't attempt to address all of the dimensions of the equity–inequity question.

Because of the coalition's regional focus and because of the interests of its community-based membership, it understood at the outset that the value of the equity analysis would rest on high-quality neighborhood-level data that would offer a more refined spatial resolution than the Orfield *Metropolitics* analysis had offered. As a result, CLF waited until the release of the Census 2000 Summary File 3 data, which included a wide variety of population and housing characteristics data at the block group and census tract levels, to begin the work.

## The Environmental Constituency

In preparation for the analyses, CLF surveyed its membership to identify potential indicators and variables to be considered in the analysis of equity. CLF is organized into topically focused workgroups that form an important structure for guiding its work; CLF encouraged these workgroups to actively participate in the creation of the Regional Equity Atlas through the contribution of data and expertise. This distribution of labor was partially a matter of scarce resources, but it also ensured that CLF members were invested in the project and the advocacy to follow. As the mapping project progressed, CLF held a series of focus groups with its members to get feedback as to the relevance and legibility of the maps that were being created.

The Audubon Society of Portland, more than any other of CLF's member organizations, recognized the potential that a fine-grained geospatial analysis of access to parks and greenspaces across the region (as opposed to simple per capita measures, such as acres per thousand residents of municipalities) could offer its advocacy work. Under the leadership of Portland Audubon, the parks and greenspaces workgroup had always been the best organized and most active of CLF's workgroups. In 2003, as the equity atlas project was beginning, Audubon and others were advocating for the revision of the regional framework plan to (1) establish levels of service, required of local jurisdictions, for a full range of park and recreational opportunities; (2) require local jurisdictions to adopt policies that would ensure the protection and restoration (if necessary) of natural areas; and (3) require the cleanup of local tributaries to the Willamette River to ensure equal access to water-based recreational activities. Just as Orfield's maps had helped to mobilize CLF in its efforts to influence the first regional framework plan, CLF's parks and greenspaces working group saw in the Regional Equity Atlas project an opportunity to mobilize and shape regional policy related to parks and greenspaces.

In April 2003, Mike Houck, a naturalist with Portland Audubon who had helped to found CLF, created a "Natural Resource Protection/Parks and Greenspaces Equity Draft Outline" for consideration by the workgroup. In this document, he suggested that the workgroup submit, for the equity atlas project, a document that would argue that natural resource protection and parks and greenspaces were equity issues with regional significance; identify one-to-three year policy or project goals that the equity atlas research would support; and outline indicators that the group wanted to see mapped and the primary questions that should be answered by the Regional Equity Atlas and its underlying research.

In fact, Audubon saw this effort as so important that it dedicated some of its own staff time to the equity atlas project. On Audubon's behalf, Jim Labbe, an Audubon conservationist who had previously volunteered with CLF, worked closely with Ken Radin, the Portland State University geographic information system (GIS) analyst on the project, to help frame the issues and assist in conceptualizing the analyses. No other coalition member organization did this. For example, Radin and Labbe developed a concept they called *park-sheds*, which was achieved through a high-resolution geospatial raster analysis (which creates a continuous mapping surface as opposed to the discontinuous surface that the choropleth mapping approach creates) of populations in relationship to parks and natural areas, to determine the percentage of the population within a quarter-mile of these assets. Labbe considered this to be an important advancement over the conventional per capita measures of park access that tend to mask the actual differences in park access at the neighborhood level. The approach also created a cartographic output by which park-deficient neighborhoods could be easily identified.

The conversion of the data to a raster format allowed CLF to present the data in a form that resonated with the Portland regional community. Portland's residents identify strongly with the neighborhood geographies within the city's boundaries, and suburban neighborhood associations and community planning organizations are also an important part of the region's civic infrastructure. Although the conversion of the data to a raster format was primarily accomplished for the purposes of geospatial access analysis, it also allowed the analysts to summarize the various data by the region's neighborhood and community planning organization boundaries, areas that are much more familiar to the general public and policymakers than, for example, census tracts. A by-product of the GIS analysis was a neighborhood summary table that provided demographic and access statistics for each of the region's neighborhoods. CLF added a normalized 1 to 4 ranking system of the neighborhoods to the summary table to allow for an easy comparison of one neighborhood to another for each of the variables.

This combination of tabular neighborhood statistics and park-shed maps provided the parks and greenspaces workgroup with the evidence it needed to confirm that the distribution of parks and natural areas was not uniform and to identify which areas were most deficient. Having been intimately involved in developing the metrics around park and greenspace access, and given his expertise with conservation policy, Labbe was the obvious choice to author the chapter on access to parks and natural areas in the Regional Equity Atlas.

## The 2006 Natural Areas Bond Measure

The Regional Equity Atlas analysis of access to parks and natural areas was one of the earlier parts of the Atlas to be completed and was available before the publication of the Atlas in 2007. As the greenspaces analysis was becoming available, Metro staff and environmental advocates began to explore the possibility of expanding on the purchase and protection of natural areas that was initiated under a $136 million regional greenspaces bond measure that had been approved by voters in 1995. Although there had been significant analysis before the first greenspace bond measure of the locations and allocations of natural assets to be acquired, no analysis had specifically focused on equity and the distribution of natural assets or their locations relative to disadvantaged populations.

Coincidentally, as the second greenspaces bond measure discussion was building, CLF had launched a significant outreach effort to publicize the soon-to-be published Regional Equity Atlas. As part of this outreach, the coalition engaged the Environmental Justice Action Group and other long-time environmental activists in a discussion about how to best use the new information that the Regional Equity Atlas analyses had generated around access to parks and natural areas. According to Labbe, it was a "no-brainer" to apply what he and CLF had learned through the atlas process to integrate equity into the regional policies that would underlie the new bond measure.

Before the release of the Regional Equity Atlas, bond measure advocates had considered including a capital grants program that, unlike the regional and municipal shares that had made up the 1995 measure and would remain the backbone of the new bond measure, could accommodate unanticipated opportunities. In other words, the new bond measure could include opportunistic purchases that were not specifically identified as projects at the outset. According to Ron Carley (then CLF board chair, now executive director of CLF), what further fueled the desire to create more flexibility in the new bond measure at the neighborhood level was the discontent that had been registered among some voters who had supported the first bond measure but hadn't seen any natural area acquisition in their neighborhoods.[1] Often, these complaints came from residents of fully built-out urban neighborhoods where natural areas were scarce. Findings from the Regional Equity Atlas confirmed that there were, in fact, park-deficient neighborhoods in the region, and that these deficiencies were not being addressed by existing greenspace policies and programs. By focusing on neighborhoods, this capital grants program component could offer community groups and local planning agencies the opportunity to begin to remedy these deficiencies.

Noting that "over one-third (36%) of the region's population inside the [urban growth boundary] lives farther than a quarter-mile linear distance from a natural area" (a number based on the Atlas analysis), on February 9, 2006, CLF sent a letter to Metro Council President David Bragdon and the Metro Council that advocated for the new bond measure and the creation of a regional opportunity fund that would address neighborhood deficiencies and stress the importance of social equity. Further, CLF suggested that the Regional Equity Atlas be used as one source of information to determine which areas were "undernatured." In its letter to Bragdon and the Metro Council, CLF also strongly advocated for the inclusion of project selection criteria that would directly address undernatured areas in low-income neighborhoods not just because they lacked natural amenities, but also because they would likely lack the resources, all things being equal, to access the proposed fund:

> We also urge you to incorporate criteria that would give added weight to projects in underserved or low-income neighborhoods (e.g., neighborhoods with high poverty or high child poverty) that also have poor access to nature. The results from the Regional Equity Atlas project indicate that, while a distinct cluster of high-poverty, nature-deficient neighborhoods exists in portions of East Portland and West Gresham, there are similar pockets found throughout the entire region. These neighborhoods often lack the organizational capacity to identify projects, prepare viable applications, or secure needed matching resources.

In its letter to Bragdon and the Metro Council, CLF also advocated for points in the proposal selection process to be awarded to projects that were linked to investments with affordable housing out of concern that park projects in low-income communities might feed displacement.

In March 2006 the Metro Council referred a $227.4 million bond measure proposal to the region's voters for the purposes of purchasing natural areas and protecting water quality and wildlife habitat. The proposed bond measure would provide $168.4 million to purchase between 3,500 and 4,500 acres of land in identified target areas that would provide regional benefits in preserving wildlife habitat, protecting water quality, enhancing trails and greenways, and connecting urban areas with nature. It would also allocate $44 million to cities, counties, and park districts to fund identified local projects to protect natural areas and water quality.

In addition, it would provide $15 million of funding for a new Nature in Neighborhoods Capital Grants Program. The idea was that nonprofit organizations, local governmental agencies, and community-based organizations could apply for matched 2-to-1 funding (two dollars of outside funding or in-kind contributions for each Metro dollar) to support neighborhood-focused projects. Projects could include such activities as the acquisition of neighborhood natural areas, trail development, habitat restoration, and interpretive displays.

CLF had advocated for income and race and ethnicity to be included in the project selection criteria. And, although project proposals for the Nature in Neighborhoods program that were to be located in low-income neighborhoods were given bonus points under the provisions of the bond measure, it did not include language targeting communities of color. Nevertheless, proponents of the bond measure understood, based on the findings of the Regional Equity Atlas, that many neighborhoods that included high proportions of racial and ethnic minorities were also areas that were low-income and park– and/or natural area–deficient; thus, the provision as written could offer a previously unavailable opportunity for these neighborhoods to acquire or enhance local natural assets and thereby improve greenspace access to both low-income communities and communities of color.

## An Unanticipated Advocate

Between 2004 and 2005, Kaiser Permanente's Program Office developed a more community-based approach to health promotion under the community health initiative of its community benefit program. Nancy Stevens, director of the community benefit program for Kaiser Permanente's Northwest Region at the time, has indicated that although at first the community health initiative lacked clarity, the confluence of the recognition of the growing obesity epidemic and the notion of community planning and access to resources developed into what was called the "healthy eating/active living" or HEAL initiative. At that time, most health professionals considered obesity to be just another disease and generally treated it as if it were a clinical condition. Kaiser Permanente had been the clinical site for the MRFIT (multiple risk factor intervention trials) out of which grew the Freedom from Fat weight-loss program, which combined nutrition education, a low-fat diet, exercise, behavioral self-management, and social support designed to serve a general population and obese adults. Yet, in spite of all the research and trials that Kaiser Permanente and others had sponsored, the obesity rate was growing at an alarming rate.

During this time, Stevens and others within Kaiser Permanente became aware of the possible connection between the accessibility of goods and services in neighborhoods and the ability for community members to live actively and eat healthfully: that accessibility, diet, and physical activity were linked. In an effort to learn more, Stevens began to network with people around the region and learned that CLF's Regional Equity Atlas endeavor was focused on a geospatial assessment of the physical access to a wide variety of community assets at the neighborhood level. Intrigued, Stevens investigated CLF's work, and what she saw was compelling enough to convince her that the Atlas was an effort worthy of Kaiser Permanente's support.

As interest in the Regional Equity Atlas effort grew, support was also building for putting the new natural areas bond measure on the November 2006 ballot. The relationship between park access and opportunities for physical activity, especially for youth, seemed clear to Stevens. Moreover, the inequitable access to these assets revealed by the Atlas research suggested to Stevens that Kaiser Permanente should support the bond measure. Stevens has acknowledged that this was a naïve notion given that Kaiser Permanente had never in its history endorsed a bond measure. Now convinced of the persuasive power of maps, Stevens took national Centers for Disease Control obesity maps with her to help make her healthy people/healthy places argument in as many parts of the organization as she could. Her arguments and evidence eventually gained the attention of Cynthia Fintner, Kaiser Permanente's president at the time. When Stevens got word that Kaiser Permanente would, in fact, go on record to support Metro's 2006 natural areas bond measure, she was, in her words, "blown out of the water!"[2]

Later, when CLF needed funding to print the Atlas, Kaiser Permanente's community benefit program contributed $23,000 for that purpose. The power that the Regional Equity Atlas had demonstrated to reveal the correlations between poverty and access and poverty and health encouraged Stevens to pursue the possibility of mapping health outcome data to better understand the relationships between place, equity, and health. She was able to explore the mapping of deidentified health record data of Kaiser Permanente members in the Portland metropolitan region at the census tract level. Kaiser Permanente's in-house GIS capacity has grown since 2006, and Stevens was able to convince Kaiser Permanente to map health outcome data that coincide with the 2010 Census for use in version 2.0 of CLF's Regional Equity Atlas, which is currently under way.

### The Bond Measure Is Approved

On November 7, 2006, voters approved the $227.4 million natural areas bond measure, which included the Nature in Neighborhoods Capital Grants Program. This program specifically focuses on the desires and needs of individual neighborhoods, not municipalities or the region as a whole. Further, the approved bond measure put projects in low-income neighborhoods at the top of the list of priorities for bonus points in the proposal evaluation process (Resolution No. 06-3672B).

## The Nature in Neighborhoods Capital Grants Program

Metro announced that it would begin to accept applications to the Nature in Neighborhoods Capital Grants Program in fall 2007. However, although agencies and greenspace advocacy groups were eager to use the funds, they soon realized that it was difficult to identify the type of innovative, community-based projects that the program hoped to inspire. Mary Rose Navarro, who has overseen the program from the beginning, believes that identifying and securing support for new capital investments in the public sector seemed daunting to some community groups. Such projects take foresight, an understanding of local dynamics, an ability to build relationships with unique organizations, and patience as the political will for a project evolves. Thus, the program has not yet awarded the full $2.25 million annually that it is authorized to award.

The program is available to any neighborhood group, and several affluent neighborhoods have benefited. However, the Atlas findings pointed to several disadvantaged neighborhoods where the Nature in Neighborhoods Capital Grants Program's dollars could have a major impact. Knowing this, Navarro has dedicated considerable time to outreach and education and has used the Regional Equity Atlas to develop an outreach strategy. The results of these targeted outreach efforts have been impressive, not only in helping to preserve and/or enhance (or "re-nature") natural assets in low-income, park-deficient neighborhoods, but in building relationships that have created multiplier effects in these communities.

In the program's first round of funding, a project was developed in an urban renewal district in Clackamas County that, according to the Regional Equity Atlas, was the region's most park-deficient area. This project involved a partnership among the local development agency, a local real estate developer, and the Clackamas Community Land Trust, a nonprofit housing provider. The developer had initially received approval to build 25 housing units on four acres of land in the urban renewal district. However, the Nature in Neighborhoods Capital Grants Program presented the renewal agency with the opportunity to create an unanticipated natural asset in the district. By permitting the developer to build an additional four units of housing on three acres instead of the original four, the development agency received an acre of land for park development. In addition, as part of the project, the Clackamas Community Land Trust negotiated a permanent affordability covenant on 10 of the 29 units.

This collaborative and innovative approach to creating a natural capital asset in a park-deficient area can be attributed, in part, to CLF's own outreach to affordable housing developers for this project. Given CLF's concern that these public investments in low-income neighborhoods, although essential to the livability of these neighborhoods, could threaten their affordability, and given that the incentives CLF had advocated for inclusion in the bond measure that would have allotted bonus points to proposals that made these linkages had not been adopted, CLF believed that such outreach was essential.

This public–private–nonprofit collaboration made Navarro realize there could be many innovative ways for communities, especially low-income communities, to procure the required match, as well as leverage the grant funding, to serve a multitude of purposes. As a result, Navarro began to target not only specific geographic areas for outreach, but also potential strategic partners, such as affordable housing providers and housing authorities, for presentations about the benefits and mechanics of the Nature in Neighborhoods Capital Grants Program. This outreach approach has resulted in a number of projects in the park-deficient areas that were identified in the Atlas.

The Nadaka Nature Park is located in the City of Gresham in one of the lowest-income and most racially and ethnically diverse neighborhoods in the region. The Audubon Society of Portland, the Trust for Public Land, the East Wilkes Neighborhood Association, St. Aidan's Episcopal Church, and the East Multnomah Soil and Water Conservation District all partnered with the City of Gresham to apply for a Nature in Neighborhoods grant in 2009 for the purchase of a two-acre expansion of the park that would provide greater visibility and public access to what has been described as a "hard-to-find" and poorly maintained public asset. In the proposal, the City of Gresham committed to developing a master plan and a natural resource management plan for the expanded 12-acre park. Metro awarded $220,000 to the city for the project, which has increased the number of children served, according to Labbe, by approximately 134 percent, and the number of people of color who can easily access the park by approximately 214 percent. Since the award, the collaboration has expanded so that now Verde (a nonprofit organization that offers green jobs workforce training to low-income youth), the Columbia Slough Watershed Council, Human Solutions (a community development corporation), the Rockwood Neighborhood Association, and Catholic Charities are involved in the project, which is integrating access to nature, green jobs, and urban agriculture into what could have been solely an environmental restoration project.

Cornelius, a city of 11,875 on the western edge of the metropolitan region, is exactly the kind of community in which the Nature in Neighborhoods program could have a major impact. It is park deficient, low income, and largely Hispanic. But community members couldn't figure out how to obtain the required capital asset to qualify for funding dollars. As long as they focused on parks, trails, or greenspaces, they were stymied; there didn't appear to be any opportunities. But Navarro encouraged them to think about the kinds of public works projects they would like to see in their community. The "re-greening" aspect of the program doesn't limit the projects to existing parks, trails, and natural areas; it can apply to any public works project. Soon discussion developed around the possibility of cleaning up, for pedestrian use, an alleyway that was covered in asphalt and lined with dumpsters. It was, according to Navarro, "as far from green as you can possibly go."[3]

The first opportunity to redevelop in the area came when the Virginia Garcia Memorial Health Clinic, a nonprofit community health care provider serving migrant farm workers, received a $12 million federal stimulus package to redevelop their entire block. Across the street is the Centro Cultural, a community center that serves the Latino community. With the idea of developing an environmentally powerful and sustainable project that could qualify for funding under the Nature in Neighborhoods grant program, a partnership was developed that included not only the Virginia Garcia Health Clinic and Centro Cultural, but also Adelante Mujeres (which provides education focused on sustainable, organic agricultural practices for the clinic's community garden), the

City of Cornelius, and the Jackson Bottom Wetland Preserve. Centro Cultural would work with Verde to create a new green jobs training program for local residents that would also perform maintenance for the alley's proposed bioswales and rain garden. In 2011 the project, which is currently under development, was awarded $322,234 with a total project cost of $1.2 million.

These are just three examples (and there are many more) of the kinds of projects made possible by the Nature in Neighborhoods Capital Grants Program that was funded by Metro's 2006 natural areas bond measure. Unlike the regional and municipal components of the bond measure, which had preidentified projects, the capital grants program has the flexibility not only to respond to unanticipated opportunities (which the other two components of the bond measure cannot), but also to respond directly to the needs of individual communities at the neighborhood level in new and creative ways. As the program builds on its outreach strategy by targeting park-deficient neighborhoods and strategic partners, the number of innovative projects and partnerships has increased, bringing it closer to its potential annual allocation.

## Concluding Observations

The environmental findings of the Regional Equity Atlas were first used by CLF and Audubon to lobby for a neighborhood-focused capital grants component to the 2006 natural areas bond measure. The strength of their arguments rested on the objectivity of the data and analyses that the university provided. Their persuasiveness can also be attributed to the intimate knowledge that Labbe, in particular, had gained in helping to develop, with the Portland State University GIS analyst, the metrics and analytical framework for the environmental aspects of the Atlas. That work and Labbe's expertise as a conservationist allowed him to credibly wear two hats (that of analyst and that of advocate) in the context of arguing for the inclusion of equity as a foundation of the 2006 bond measure.

The Atlas analyses found an equity dimension to parks access in that many of the most park-deficient areas were in low-income neighborhoods populated by ethnic and racial minorities. With the Regional Equity Atlas initiative, CLF, the Audubon Society of Portland, and others were able to effectively demonstrate that social equity (focused on income and race and ethnicity), especially at the neighborhood level, was at play and could be verified, located, and measured. As a result, these equity advocates (with the help of Metro Councilor Robert Liberty, who was also a founder of CLF and former executive director of 1000 Friends of Oregon) were successful in getting equity language relating to income incorporated into the bond measure, explicitly establishing equity in Metro's regional greenspace policy. The inclusion of this language was a huge victory for equity advocates. And, although language relating to race and ethnicity was not included in the bond measure, equity advocates could assume, because of the geographic intersections between low-income neighborhoods and communities of color that the Atlas revealed, the priority given to proposals in low-income neighborhoods would affect racial and ethnic minority communities as well.

The other key concern shared by CLF and Audubon—that the affordability of neighborhoods receiving grant dollars could, as a result, be threatened—was not addressed by the bond measure. However, CLF and Audubon have taken on making those linkages a priority at the outreach, proposal development, and even postproposal stages for projects.

Further, the Regional Equity Atlas neighborhood-level analyses relating to parks and greenspaces had some unanticipated impacts. The Atlas effort was instrumental in persuading a nationally influential health institution, Kaiser Permanente, to take the unprecedented step of publicly supporting a bond measure that it saw as promoting its community health agenda.

On the environmental organizational front, although Houck and other greenspace advocates were quite certain, even before the analyses began, that there was an equity dimension to park and natural area access, they did not anticipate the extent to which the findings of the Atlas would influence Portland Audubon's own strategic planning. The Audubon Society of Portland has been located in the relatively affluent Portland westside since 1949. The Atlas analyses made clear not only the intersection between park deficiencies, low-income populations, and ethnic and racial minorities, but also the large number of underserved children living in Portland's outer eastside and Gresham. In July 2010 the Audubon Society of Portland opened its first eastside satellite office at the Leach Botanical Garden, which is within walking distance of two of the region's lowest-income neighborhoods, Powellhurst-Gilbert and Lents. Audubon currently offers bird and natural history classes, walks, presentations, and summer camp for youth out of that facility. Labbe has said that not only was this the right thing to do, but it was also strategic given Audubon's aging membership:

> As the Equity Atlas demonstrated, we're becoming a much more diverse community. Particularly in the East Metro community you see that significantly in school-aged children. If you compare school-aged children of people of color versus the adult population, I think it's like about 25 percent adult population in Multnomah County but it's like upwards to 40–45 percent children. So, you know, the writing's on the wall. I'm really glad that Audubon is being strategic about thinking about the next hundred years and how we're going to be relevant.[4]

## NOTES

1. Ron Carley (board chair, Coalition for a Livable Future), interview with the author, January 27, 2012.

2. Nancy Stevens (director, community benefit program, northwest region, Kaiser Permanente), interview with the author, September 29, 2011.

3. Mary Rose Navarro (Nature in Neighborhoods Capital Grants Program), interview with the author, October 11, 2011.

4. Jim Labbe (conservationist, The Audubon Society of Portland), personal interview, October 7, 2011.

## REFERENCES

Coalition for a Livable Future. 2007. *Regional Equity Atlas: Metropolitan Portland's Geography of Opportunity.* Portland, OR: Coalition for a Livable Future. http://www.equityatlas.org.

Walljasper, Jay. 1999. *Working for a Just and Sustainable Region: An Uncommon Alliance with a Common Vision.* Portland, OR: Coalition for a Livable Future.

ESSAY

# A Solvable Problem Hidden in Plain View

## *When Monitoring the Right Data Makes All the Difference*

Hedy N. Chang

Research confirms what we know from common sense: missing too much school makes it harder to succeed in school. Beginning in kindergarten and even preschool, chronic absence (missing 10 percent or more of school for any reason, including excused as well as unexcused absences) can predict lower third-grade reading scores, especially for children living in poverty who experience more than one year of chronic absence (Chang and Romero 2008; Connolly and Olson 2012). By middle and high school, chronic absence is a proven early warning sign of dropping out of high school for all students regardless of their socioeconomic status (Spradlin et al. 2012). If too many students have poor attendance, the classroom churn can slow down learning for everyone (Musser 2011).

What is much less well-known is that chronic absence doesn't affect just a handful of children. An estimated 7.5 million students (1 of 10) in the United States miss so much school that they are academically at risk. In some communities and schools, more than one of four children are chronically absent.

Yet many schools and communities don't know how many and which students are chronically absent. Most schools monitor average daily attendance and truancy rates, but both of these indicators can mask high levels of chronic absence. Consequently, schools and communities are missing out on the opportunity to use an easy-to-understand, readily available attendance measure that could help them engage in prevention and intervention strategies that could significantly improve outcomes, especially for our most vulnerable children.

The good news is that chronic absence can be reduced if schools, community, and parents and families work together to monitor the data, identify and address common barriers to getting to school, and nurture a local culture of regular attendance. This essay describes why paying attention to chronic absence matters. It also discusses how community-level practice, policy, and evaluation can make a recognizable difference and why neighborhood or other place-based initiatives should make monitoring and addressing chronic absence an integral component of their work.

## What Led to the Discovery of Chronic Absence?

The Annie E. Casey Foundation's deep commitment to a data-driven approach to improving outcomes for children, families, and neighborhoods led to the discovery of chronic absence as a key challenge to be addressed. The foundation recognized that reading proficiently by the end of third grade is critical to ensuring children have the chance to be economically successful when they reach adulthood. To inform their emerging approach to advancing educational outcomes for children in the Making Connections neighborhood initiative, Annie E. Casey Foundation Executive Vice President Ralph Smith asked me to find research and best practices that could answer two questions: Does missing too much school in the early grades contribute to low levels of third-grade reading proficiency, especially among young children in low-income neighborhoods? What are promising or proven practices for improving attendance?

The key to answering these questions lay in finding and mining local and national data sources that could deepen our understanding of the prevalence, impact, and potential causes of chronic absence. The National Center for Children and Poverty painted a national picture using attendance data gathered through the Early Childhood Longitudinal Study–Kindergarten Cohort. Equally important, a rich collaboration with the Urban Institute, the National Neighborhood Indicators Partnership (NNIP), the National Center for School Engagement, and Metis Associates supported gathering and analyzing attendance data from nine localities. Documenting for the first time the prevalence, consequences, and potential causes of chronic absence, the report *Present, Engaged, and Accounted For: The Critical Importance of Addressing Chronic Absence in* *the Early Grades* (Chang and Romero 2008), is based on the results of these data analyses as well as an in-depth literature review and scan of best practice. Ultimately the insights gained led to the development and launch of Attendance Works, a national initiative aimed at addressing this hidden but solvable crisis.

## What Is Chronic Absence?

In the broadest sense, chronic absence refers to missing so much school for any reason that a student is academically at risk. How states and local communities define chronic absence varies significantly. Maryland, for example, tracks the number of students who have missed 20 or more days; Georgia monitors students who have missed 15 or more days.

Attendance Works recommends defining chronic absence as missing 10 percent of school for any reason. Research shows an association between missing this much school and lower academic performance. Chronic absence is especially problematic for the long-term academic performance of low-income students who depend more on school for opportunities to learn. Because they are more likely to face systemic barriers to getting to school, low-income children, many of whom are children of color, have higher levels of chronic absence (Chang and Romero 2008). By middle and high school, high levels of chronic absence are associated with dropping out of school for students of all socioeconomic backgrounds (Spradlin 2012).

In addition, using 10 percent versus a specified number of days promotes early identification of students. Schools can use missing 10 percent or only two or three days at the end of the first month of school to trigger intervention rather than waiting for a student to miss 20 days. This definition also allows for better detection of attendance problems among highly mobile students, who often move too frequently to ever accumulate 20 days of absence in a single school or district but are likely to show up as missing 10 percent even with a short tenure. Finally, using a 10 percent measure allows data to be compared across districts with different lengths of school year.

To advance the work, Attendance Works recommends adopting 10 percent as a common definition across districts and states. Using this definition would help to create a shared understanding of what the issue is and why it matters across all stakeholders: educators, parents, community partners, policymakers, and the media. It would also promote better comparisons of what works across communities and make integrating chronic absence into tools for schools and communities much simpler, whether the tools are student information systems or school district report cards.

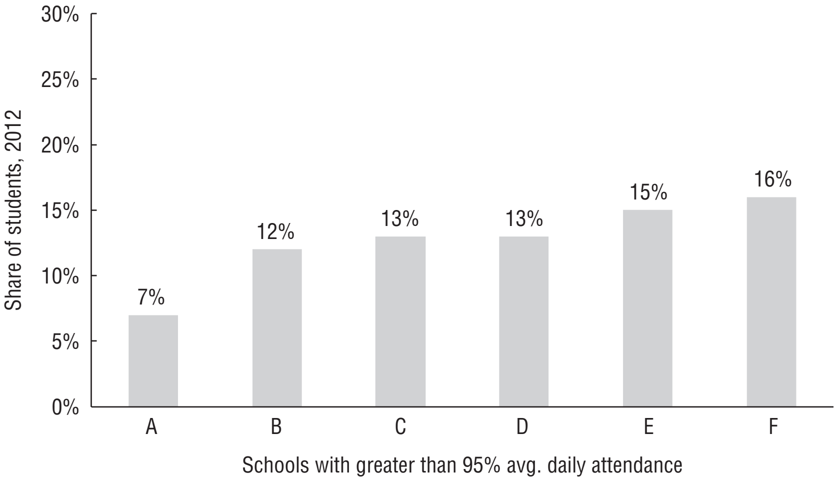
## Why Is Chronic Absence Overlooked?

A significant challenge is the lack of awareness among schools and communities about the extent to which chronic absence is a problem. Why does chronic absence go undetected even though most teachers take roll every day? First, chronic absence can be hard to notice if communities rely only on teacher observation. Especially with increasingly large classes, teachers can easily overlook a child who is chronically absent, especially if absences are sporadic, occurring once every few weeks rather than all in a row. Fortunately, most districts can now take advantage of their electronic data systems to track and monitor attendance, though often some extra steps are needed to ensure their systems calculate chronic absence rates and generate a list of the students who are at risk due to missing too much school.

Second, many schools and districts may not realize they have a significant chronic absence problem because they use the average daily attendance rate as the yardstick. Average daily attendance refers to the percentage of students who typically show up every day. Unfortunately, even an average daily attendance rate of 95 percent can mask chronic absence. Consider a school of 200 students. Ten students aren't in their seats each day, leaving the school with 95 percent average daily attendance (a level typically seen as good). Over the course of a year, those 10 students aren't the same or they would be disenrolled. The question to answer is whether the 10 empty seats reflect most of the students missing a few days or whether they are typically the seats of a small but still significant minority of students who are each missing nearly a month of school (i.e., are chronically absent) over the course of the school year.

Consider these data from Oakland, California (figure 6.2.1), which show the range in chronic absence across elementary schools, all of which had average daily attendance of 95 percent. Although chronic absence was only 7 percent in school A, it was more than twice that level in school F.

**Figure 6.2.1.** Chronic Absenteeism Rates Vary in Oakland, California Schools with High Average Daily Attendance

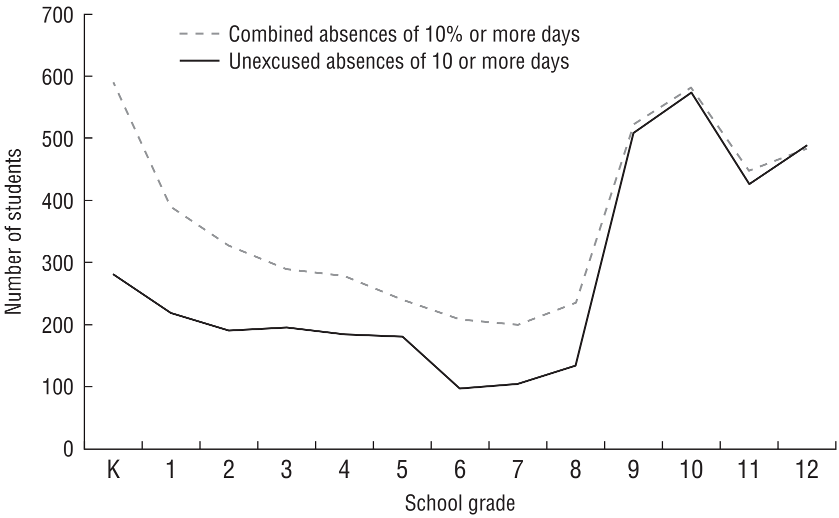



A review of data from three districts revealed that if schools had 98 percent average daily attendance, they typically had relatively low levels of chronic absence of 5 percent or less. If a school had 93 percent average attendance, however, often chronic absence affected 20 percent or more of the students. Based on average daily attendance alone, it was difficult to determine whether chronic absence was a problem for large numbers of children if a school had 95 percent average daily attendance (Bruner, Discher, and Chang 2011).

Third, many schools and communities may not realize that tracking truancy is different from monitoring chronic absence. Often, the terms truancy and chronic absence are used interchangeably, but they don't mean the same thing. Truancy generally refers to unexcused absences, although the precise definition of what constitutes truancy is determined by each state based on No Child Left Behind. For example, in Utah, students are counted as truant if they have 10 unexcused absences, but in Maryland the trigger is missing 20 percent of the school year due to unexcused absences.

Truancy overlooks, however, when children miss a lot of school but the absences are excused. Particularly when children are young, they can miss a lot of school due to excused absences. Consider figure 6.2.2. It shows

**Figure 6.2.2.** Truancy Alone Underestimates Chronic Absentee Problem in San Francisco Unified School District, 2010/2011




*Note:* Unexcused absences only include those through May 16, 2011.

that using chronic absence to identify students at risk due to poor attendance identifies significantly more children than looking only at those who have accumulated large numbers of unexcused absences. The difference is especially noticeable in kindergarten. Although 5- and 6-year-olds generally are not missing school without the knowledge of an adult, they can still miss so much school that their academic progress is affected.

For districts and communities to fully understand attendance patterns and challenges among their students and schools, they need to calculate chronic absence in addition to the more typically calculated measures of average daily attendance and truancy. Each measure offers different insights into what is happening around attendance. Average daily attendance paints a picture of how many students show up on any given day, but chronic absence reveals whether a significant number of students are missing so much school that they are at risk. Truancy data help families, schools, and communities identify if many students are missing school without permission.

Chronic absence is arguably, however, even more overlooked as an issue in preschool and child care programs for a whole variety of reasons.

First, preschool and child care programs are part of a highly fragmented system with multiple funding sources and no common data system. Second, attendance is not always collected and if it is, it is typically for funding purposes, not necessarily for identifying students at risk. Moreover, although some larger programs like Head Start or state preschools typically have some form of attendance data that can be manipulated to calculate chronic absence, those programs might not always have staff with the data sophistication to conduct the analysis. Finally, preschool programs, like school districts, are also confused by average daily attendance. Many may not realize the 85 percent average daily attendance required by Head Start can mask extremely high levels of chronic absence. Pioneering work by the Baltimore Education Research Consortium and the Chicago Research Consortium now shows that chronic absence starting in preschool can predict lower elementary school performance, especially if chronic absence occurs for multiple years. The levels of chronic absence can be extremely high, affecting 40 percent of publicly funded preschool programs (Connolly and Olson 2012; Ehrlich et al. 2013). Few preschool programs, however, have ever examined if chronic absence is a problem.

Fourth, even if districts or preschools have taken the step of calculating chronic absence, key community stakeholders may still be unaware that chronic absence is a problem because they have never seen the data. Sometimes districts do not share the data because they are concerned about protecting confidentiality. They may not realize that sharing aggregate data on overall levels of chronic absence by grade, school, or student subpopulation is not a violation of either FERPA or HIPAA as long as the number of students included is large enough to avoid attributing the data to an individual student. Districts may also be concerned about releasing data because they fear the data will be used to cast blame rather than create the conditions for community stakeholders to partner effectively with schools to address barriers to attendance.

To change this situation, schools (including preschools) and communities need to establish forums in which chronic absence data can be shared so community stakeholders can work together to analyze what the data mean and determine their implications for allocating and leveraging local resources that can help students get to school. Ensuring students get to class every day is not a problem for schools or preschools alone to solve, although school districts are a key first stop on the path forward as they are the ones with the data essential to triggering and informing action.

## What Contributes to Chronic Absence?

Chronic absence reflects the degree to which schools and preschools, communities, and families adequately address the needs of students. Attendance is higher when schools and preschools provide a rich, engaging learning experience; have stable, experienced, and skilled teachers; and actively engage parents in their children's education. Chronic absence decreases when educational institutions and communities actively communicate the importance of going to school regularly to all students and their parents starting with the entry to school, and reach out to families when their children begin to show patterns of excessive absence. Attendance suffers when families struggle to keep up with the routine of school and lack reliable transportation; work long hours in poorly paid jobs with little flexibility; live in unstable and unaffordable housing; have inadequate health care or suffer from a prevalence of chronic disease; and experience escalating community violence.

Taking time to unpack why students miss school in the first place is essential to developing effective solutions. Attendance Works has found it helpful to classify the reasons students miss in terms of three broad categories:

- *Myths.* A number of common and pervasive myths about attendance make it less likely that going to school every day is made a top priority. Often, good attendance is seen as a matter of complying with rules rather than a matter of providing children more and better opportunities to learn. Consequently, missing school is only seen as a problem if a child skips school without permission. Often families and educators do not realize that too many absences, even if they are excused, can quickly add up to so many that they hinder learning and just missing two or three days every month is a problem. Many of them do not recognize that poor attendance as early as preschool and kindergarten can have a detrimental impact on their child's ability to succeed in school.
- *Barriers.* Many students can't get to school as a result of chronic health conditions or inadequate access to medical, mental health, or dental care; unstable housing; unreliable transportation; or a lack of effective family and community support and service delivery. These barriers are especially critical for children involved in foster care or the juvenile justice system. An analysis by the Uni-

versity of Utah, for example, found that homeless students were 2.5 times more likely to be chronically absent. Among these homeless students, some living conditions (like living in a shelter) were associated with even higher absenteeism than others (such as living with another family) (Spradlin 2012).

- *Aversion.* Sometimes poor attendance occurs when students avoid going to school in response to, for example, bullying, an unhealthy school climate, punitive unfair disciplinary practices, or ineffective instruction. Analyzing chronic absence data by classroom can help to reveal if the problem is schoolwide or concentrated in particular classrooms. In some cases, the aversion isn't just on the part of a student. Poor attendance could reflect the fact that parents had a negative experience with school and they have not been assured that their child's experience will be different.

What keeps a particular student or group of students from getting to school or preschool will vary significantly by student, school, and community. But keeping these three categories in mind can help identify the biggest challenges for the largest numbers of students so appropriate programmatic interventions can be put in place.

A key role that community or neighborhood initiatives can play is helping schools and preschools to gather and analyze data so they have a deeper understanding of local attendance barriers. The size and scale of the chronic absenteeism problem can offer clues about the nature of the attendance challenges. Students and families with the most severe levels of absence often face multiple barriers to getting to class. If only a small number of students are chronically absent, then issues are more likely to be individual in nature. When chronic absence affects large numbers of students in a particular school or neighborhood, it is often an indication of more systemic challenges. Combining the data from schools with data from other sectors, such as health, housing, and transportation, can shed additional light on the root causes of the problem. Attendance Works (2010b) has developed guidance for drawing on a combination of qualitative and quantitative sources of information that can identify the factors that contribute to chronic absence in schools.

Students and their families are especially critical sources of information about barriers. Communities can solicit their insights through a variety of techniques ranging from focus groups and surveys to looking for patterns in data collected from families by case managers. The Baltimore Education

Research Consortium used the data collection process to engage parents of preschoolers in the work. They provided parents with cameras, for example, so they could take pictures of what makes it difficult to get to school or preschool. One parent, for example, shared a photo of a daunting and wide road without a crosswalk.

Geomapping is another essential tool because it allows for examining patterns by neighborhood. In Oakland, for example, the Urban Strategies Council, a founding NNIP member, used this technique to reveal that chronic absence in the early grades is concentrated in West Oakland and some parts of East Oakland (see figure 6.2.3). Additional data sources revealed that these are also neighborhoods with high levels of poverty and crime, as well as significant health concerns, such as elevated rates of asthma.

## What Turns Chronic Absence Around?

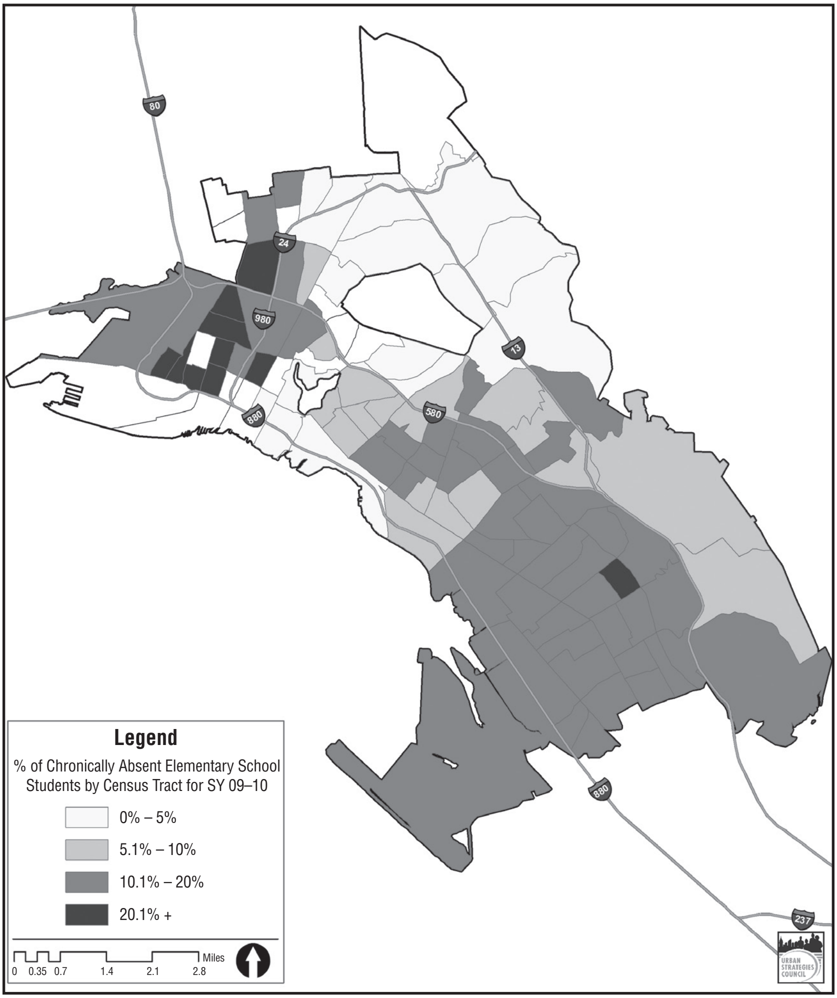
The good news is that schools, working in partnership with families and community agencies, can turn chronic absence around. One important approach to identifying what works is to seek positive outliers. Find the school communities with high levels of poverty but low levels of chronic absence and then conduct research to identify what is happening to improve student attendance. Such research not only can help illuminate best practices but also identify administrators, teachers, attendance clerks, and community partners who can serve as inspiring examples of what is possible.

Through observing such positive outliers, reviewing available research, and working with districts over time, Attendance Works has found that two levels of strategies are essential to reducing chronic absence and sustaining progress. First, at the school-site level, chronic absence can be decreased substantially when school communities[1] use five specific strategies (see figure 6.2.4) to nurture a culture and a habit of attendance while also identifying and addressing barriers to getting to school. Second, schools are more likely to adopt and sustain these five practices when districts and their community partners have put in place four supporting key actions (described later in figure 6.2.5).

The school-level strategies, including the role of data, are listed below; the higher-level strategies are described after this list.

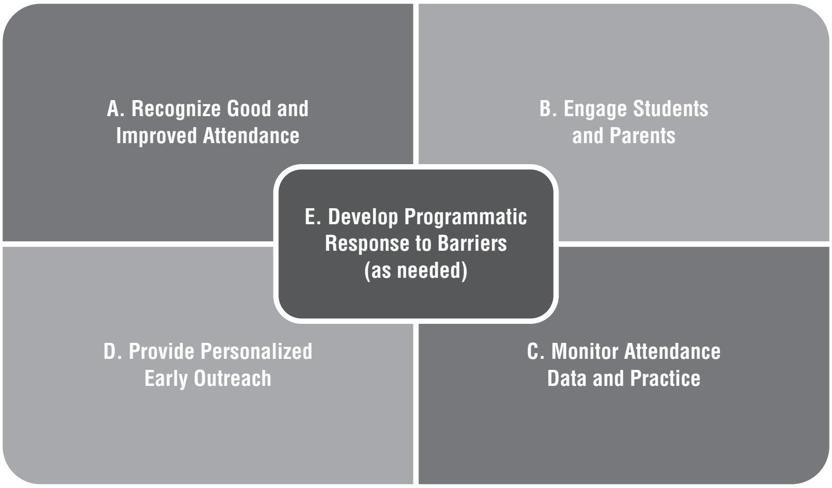
A. *Recognize good and improved attendance.* School communities can send a clear message that going to school every day is a priority by

**Figure 6.2.3.** Share of Elementary Students Chronically Absent in the Oakland Unified School District, 2009/2010




*Sources:* 2000 Census Tract, Esri, Oakland Unified School District as produced by the Urban Strategies Council on March 4, 2011.
*Note:* Chronic Absenteeism is defined by missing more than 10% of enrolled school days in a year. Students enrolled less than 45 days are not included in these data.
A color version of the map is available at http://prezi.com/c6zuasjran10/oakland-chronic-absenteeism/

**Figure 6.2.4.** Strategies for School Communities to Reduce Chronic Absence




providing regular recognition and rewards to students and families who have good and improved attendance. Keep in mind the goal is not to focus on perfect attendance, because the children who struggle the most will soon be left out of such awards. Attendance Works (2010a) has created attendance incentive guidelines and examples of the best ways to recognize good and improved attendance. This strategy can also help improve the accuracy of attendance data, because the students themselves are likely to help ensure teachers are aware of who is and isn't in class!

B. *Engage students and parents.* Attendance improves when a school community offers a warm and welcoming environment that engages students and families in the life of the school, a competent staff, and enriched learning opportunities. A key component of the engagement is helping families learn about the positive impact of good attendance and the negative impacts of chronic absenteeism on realizing their hopes and dreams for their children. Parents may not realize that even excused absences can, if they accumulate, cause their child to fall behind and that whether they build the habit of attendance in the early grades could affect their children's chances of graduating from high school.

C. *Provide personalized early outreach.* Perhaps the most critical strategy is using the data to trigger early caring outreach to families and students who are already missing too many days of school to help the student return to school. Regular review of attendance data will reveal which children are chronically absent. Once the students at risk are identified, the outreach is best carried out by an adult with a strong relationship with the family, but the person who engages in the outreach can vary.

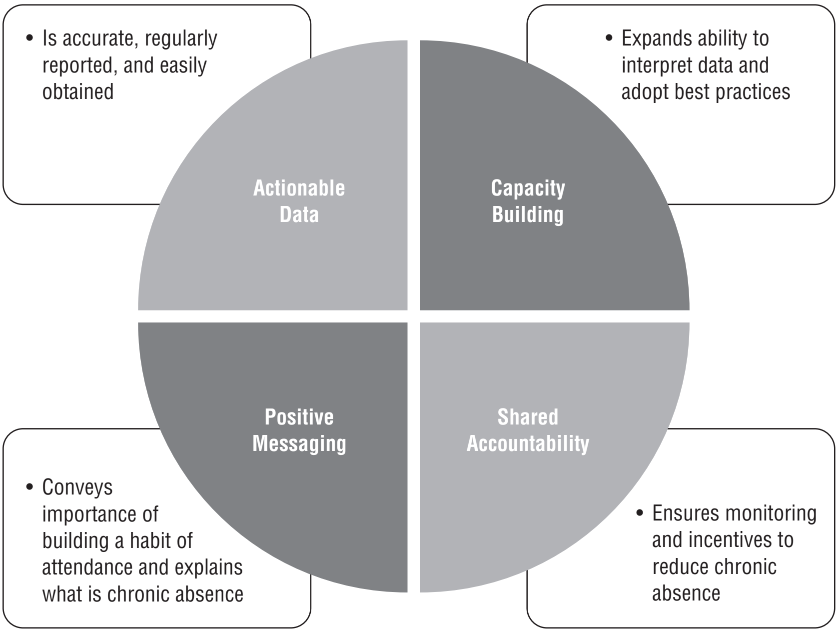
In New York City, for example, a corps of Success Mentors works directly with chronically absent students. Each mentor is assigned to chronically absent students identified by the school's principal and attendance team. If the students don't turn up for school, the mentor calls them to find out where they are. If a student is struggling with class work or social dynamics, the mentor is there to help. Mentors come from a variety of sources including national service members, social work interns, school faculty, and even seniors assigned to help incoming freshman.

Regardless of who conducts the outreach, the goal is for a caring person to connect with a student and his or her family to encourage good attendance and express concern when absences occur. Outreach is essential for learning the barriers to attendance for an individual student and what resources, such as food, health, shelter, transportation or other resources, would help improve attendance.

D. *Monitor attendance data and practice.* Each school should have a team in place that meets regularly to review the school's attendance data and coordinates efforts to reduce chronic absence. This team could be devoted exclusively to attendance or include attendance as one of other responsibilities and functions. Attendance Works (2012) has created a school self-assessment to aid the team engaged in this monitoring to examine current strengths and gaps so they can determine what needs to occur to ensure these strategies are fully in place so the desired progress can be achieved.

E. *Develop programmatic responses to systemic barriers.* As discussed above, if large numbers of students are affected by chronic absence, then it is likely that some type of systemic barrier or barriers are at play. Once the barriers are identified, data can be used to garner help in addressing the challenges, whether that involves establishing uniform closets, improved access to health care, walking school buses, tutoring, mentoring, or other types of activities.

**Figure 6.2.5.** Key Community Capacities Needed to Support Chronic Absenteeism Interventions




The five school-level strategies described above are much more likely to be adopted and sustained when districts and their community partners have put in place the key ingredients shown in figure 6.2.5 and described here.

*Actionable data.* Taking action requires having accurate, easily accessible, up-to-date data on which and how many students are chronically absent, preferably by school and grade. Ideally such information about which students are chronically absent should be available and reviewed monthly. Data on overall levels by school, grade, and subpopulation might be examined less frequently, possibly at the end of each quarter or semester.

Whether actionable data exist depends on whether districts have the policies, attendance practice, and student information infrastructure to provide such data. Community partners, however, can help with calling for such data to be available and, depending on who they are, provide or raise resources to help develop this infrastructure and easy-to-understand data reports. Attendance Works offers a free Excel-based district and school attendance tracking tool to help interested districts calculate which students are chronically absent and generate charts on levels of chronic absence, or at least determine what they might want to build into their existing data dashboards and reports.[2]

*Capacity building.* Building the skills and knowledge of school staff and community partners to understand the nature of chronic absence (including how it is different from truancy) and what are best practices to ensuring daily attendance is essential. Analysis of local data can highlight schools where chronic absence is a problem and how many children are affected. Many school personnel don't know the difference between chronic absence and truancy or don't recognize the importance of bolstering our investments in prevention and early intervention before resorting to more expensive legal intervention strategies. Consequently, districts and key community stakeholders need to determine how they can best build capacity for implementation. Often, this capacity building requires integrating a more explicit focus on chronic absence into existing professional development.

*Positive messaging.* Data and staff expertise alone are not enough to build the community's understanding of the importance of attending school regularly. The goal of positive messaging is to help parents and students realize that daily attendance is key to reaching their dreams of a successful future in school and in life. It is an intentional shift from using the threat of fines or going to court to compel attendance to starting with an emphasis on encouraging families to take advantage of the opportunity for their children to learn in the classroom. Positive messaging takes advantage of the possibility of creating change by debunking the myths that get in the way of students going to school every day and drawing on the hopes that all families carry for the next generation.

Fortunately, everyone in a community—from the superintendent and the mayor to afterschool care providers or businesses and faith-based organizations—can play a role in sending the message. Districts and communities can produce parent flyers, public service announcements, and other materials that convey why going to school every day matters. They can also take advantage of back-to-school time to establish norms of daily attendance that will lay the foundation for the remainder of the school year. Together with America's Promise, the Campaign for Grade Level Reading, Civic Enterprise, and Points of Light, Attendance Works is encouraging communities throughout the United States to make September Attendance Awareness Month.[3]

*Shared accountability.* For attention and action to be sustained, chronic absence needs to be built into ongoing accountability systems used by district to measure progress and identify where additional support is needed to improve student performance. For example, schools should be required to examine the extent to which chronic absence is a problem and if it is, describe in their school improvement plans how they will improve student attendance, especially among the most vulnerable populations. Schools should know whether students are missing so much school they are unable to benefit from investments in improved teaching and curriculum or if they are struggling academically even though they show up to school every day.

Chronic absence can also be built into contracts with community-based organizations offering services to students at schools. For example, the Family League of Baltimore requires after-school programs to serve students with poor attendance and monitor whether their services are helping to reduce chronic absence. Evaluations of local after-school programs also show that participation in a high-quality after-school program improves the attendance of students with a past history of chronic absence (Olson, Connolly, and Kommajesula 2013; Traill, Brohawn, and Caruso 2013).

## What Are the Implications of Chronic Absence for Neighborhood Initiatives?

Chronic school absence is a natural issue for neighborhood initiatives to tackle and address. It is an easy-to-understand measure that reflects the conditions of families, neighborhoods, and schools and can be reduced through successful collaboration. Easily monitored to detect changes over time, it can help communities and schools demonstrate that programmatic interventions are having a positive impact. It is a leading indicator that can help initiatives determine if they are on the right track because attendance data are collected every day. Chronic absence data are available well before many other measures that are collected only periodically through the year.

Neighborhood initiatives can play a significant role in helping expand attention and action around the issue of chronic absence if they work in partnership with school districts to perform the following tasks:

1. Support regularly calculating and reporting on levels of chronic absence, defined as missing 10 percent of school for any reason, by

school, district, grade, and neighborhood and, if appropriate, by school feeder patterns.

2. Ensure attendance data are included in longitudinal data systems that promote sharing and tracking of outcomes across agencies.
3. Triangulate chronic absence data with data on health, social, economic, and other community conditions that could shed light on key attendance barriers facing students overall as well as those in certain neighborhoods or particular ethnic, linguistic, or economically vulnerable populations.
4. Educate local stakeholders about how they can share data on attendance while respecting concerns about confidentiality. Explain how confidentiality does not need to be a barrier, especially for looking at data at a systems level—though confidentiality should be ensured when an individual intervention is needed.
5. Convene key stakeholders to help interpret the data and identify the resources that can be leveraged to overcome barriers to attendance.
6. Encourage the use of chronic absence as a common measure to be addressed and monitored across multiple local initiatives, from Promise Neighborhoods to local grade-level reading campaigns.
7. Include chronic absence in the evaluations of local programs and publicize the results in order to expand knowledge of what interventions help to improve attendance and for which populations of students.

## Conclusion

Chronic absence is not created or solved by schools alone. This essay has suggested practical ways for schools, districts, communities, and neighborhood initiatives to tackle this issue. Whatever the level, the strategies begin with understanding the size and nature of the problem through education and other contextual data at various levels: the individual student, the school, the neighborhood, and districtwide. By pairing the analysis with proven interventions, communities can create the conditions that will ensure a next generation of children has an equal opportunity to learn and succeed.

## NOTES

1. A school community is defined as the school staff, community organizations, volunteers, students, and families associated with a school site.

2. For tools on calculating rates of chronic absence, see http://www.attendanceworks.org/tools/tools-for-calculating-chronic-absence/.

3. For more information on September Attendance month, see http://www.attendanceworks.org/attendancemonth/.

## REFERENCES

Attendance Works. 2010a. "Establishing School-wide Attendance Incentives." http://www.attendanceworks.org/wordpress/wp-content/uploads/2010/05/incentives.pdf.

———. 2010b. "Understanding the Factors Contributing to Chronic Absence in Your School." http://www.attendanceworks.org/wordpress/wp-content/uploads/2010/05/ID-Cont-Factors-DEC-2010-.pdf.

———. 2012. "Does Attendance Really Count in Our School?" http://www.attendanceworks.org/wordpress/wp-content/uploads/2012/06/School-Self-Assessment-Team-Rev-June-2012.pdf

Bruner, Charles, Anne Discher, and Hedy Chang. 2011. *Chronic Elementary Absenteeism: A Problem Hidden in Plain Sight.* Des Moines, IA: Child and Family Policy Center and Attendance Works.

Chang, Hedy N., and Mariajosé Romero. 2008. *Present, Engaged, and Accounted For: The Critical Importance of Addressing Chronic Absence in the Early Grades.* New York: National Center for Children in Poverty, Columbia University.

Connolly, Faith, and Linda S. Olson. 2012. *Early Elementary Performance and Attendance in Baltimore City Schools' Pre-Kindergarten and Kindergarten.* Baltimore, MD: Baltimore Education Research Consortium.

Ehrlich, Stacy B., Julia A. Gwynne, Amber Stitziel Pareja, and Elaine M. Allensworth. 2013. *Preschool Attendance in Chicago Public Schools: Relationships with Learning Outcomes and Reasons for Absences: Research Summary*. Chicago: University of Chicago Consortium on Chicago School Research.

Musser, Martha. 2011. *Taking Attendance Seriously: How School Absences Undermine Student and School Performance in New York City*. New York: Campaign for Fiscal Equity.

Olson, Linda S., Faith Connolly, and Alok H. Kommajesula. 2013. "Family League 2011–12 Out of School Time Programs in Baltimore City." Baltimore, MD: Baltimore Education Research Consortium.

Spradlin, Terry. 2012. *Chronic Absenteeism.* Research brief. Salt Lake City: Utah Education Policy Center, University of Utah.

Spradlin, Terry, Katherine Cierniak, Dingjing Shi, and Minge Chen. 2012. "Attendance and Chronic Absenteeism in Indiana: The Impact on Student Achievement." *Education Policy Brief* 10 (2). Bloomington: Indiana University Center for Evaluation and Education Policy.

Traill, Saskia, Katie Brohawn, and Chris Caruso. 2013. *More and Better Learning: Year One Report on ExpandED Schools.* New York: TASC.

# 7

# Advances in Analytic Methods for Neighborhood Data

Previous chapters in this volume document the progress that has been made in building neighborhood data systems and using data to drive policy and action in neighborhoods and metropolitan areas. At the same time, community-based and academic researchers are paying increased attention to how these data should be analyzed in order to provide accurate and meaningful answers to the challenging questions that face the fields of neighborhood and community change. The analytic methods applicable to neighborhood data are advancing well beyond the simple mapping and tracking of indicators for reporting or planning purposes. In this chapter, we examine a selection of recent methodological advances that relate to research on neighborhood and community change.

This chapter is organized as follows. First, we lay out some of the purposes of neighborhood data analysis and the multiple levels of analysis that are of interest to communities and researchers. This brief overview is followed by a review of selected methodologies and tools that represent promising approaches to addressing the range of applications common in the field. The techniques covered in the chapter are by necessity a subset of the vast amount of important analytic research that is taking place on neighborhoods. Even so, the chapter is not able to cover these techniques in depth. Instead, it identifies key issues and refers the reader to primary sources for details on how these methodologies can be implemented. Following the chapter are three essays that provide a cogent analysis of several analytic challenges and some of the cutting edge methods that are being developed to address them.

## Background and Framework

How does research design and analysis fit into the community indicators and community change agenda? Although much has been accomplished on the ground with sets of indicators displayed in maps and graphs, deeper and more complex analysis is required to inform policy and to add to the knowledge base about community development and change. In particular, more sophisticated analysis tools are needed to assess the quality of neighborhood measures, uncover relationships among indicators, describe places along multiple dimensions, and discover patterns of change across time and space. Moreover, efforts to estimate the effects of neighborhoods on residents or to evaluate the impact of community change initiatives pose challenges of causal attribution that need to be addressed by careful research design and analysis. In addition, the plethora of geospatial data and growing demand for information on the dynamics of neighborhood change require greater attention to the precision of community measurement and spatial processes than may have previously been assumed.

There is no one discipline or source to turn to for research design and analysis methods that are potentially useful for community indicators. In fact, the individuals contributing to this methodology are investigating a wide variety of topics and come from many backgrounds, including economics, sociology, epidemiology, geography, public policy, urban affairs, community psychology, social work, and others. Their studies have in common the use of data collected on small areas such as neighborhoods and rigorous attempts to manipulate the data in ways that address important questions about community context and change. Although not solely directed at the analysis of community indicators, there are elements of their techniques that can be used and applied.

Before we describe selected methodological advances, it is necessary to acknowledge that the choice of analytic tools and approaches must be guided by the purpose and focus of the neighborhood indicators study or project. The particular methodological difficulties that face data analysts are driven, in part, by the research questions they are asking and the phenomena they are studying. Below we offer two dimensions along which to think about analytic needs: the purpose of the analysis and the main concepts of interest.

## Purpose of Analysis

The primary purpose of a study determines what elements of research design and statistical analysis need to be emphasized. Or put another way, depending on the main purpose of an analysis, there are particular challenges to obtaining precise, unbiased, and dependable findings. The methodological advances discussed in this chapter produce findings that are as valid as possible with respect to the purpose of the analysis. This is not to say that other issues are irrelevant, but since perfection in research is seldom possible to achieve, practical considerations often dictate giving priority to some concerns over others. Consistent with social science research in general, the purposes of neighborhood indicators studies include *description, classification,* and *explanation and prediction.*

Description is a common purpose of neighborhood indicators analysis. Sometimes description is an end in itself, but accurate description can also be the first step in building a more advanced model for explanation or prediction. In descriptive studies there is an emphasis on making precise and unbiased estimates of the level or range of neighborhood attributes or conditions. Neighborhood studies can be plagued in this regard by problems of data sparseness, incomplete data, or ambiguity of neighborhood definitions and boundaries. These challenges and some promising approaches to address them are covered in this chapter.

A second purpose of neighborhood data analysis is classification of communities or conceptual domains. Classification may involve grouping places into categories or types based on similarities and differences among cases along a variety of dimensions. In addition to requiring the precision of measurement mentioned above for descriptive studies, classification requires a method for uncovering patterns of similarities and differences to be applied to the data. A valid classification is one in which cases can be assigned to a category with the least ambiguity along reliable dimensions of interest and in which the classification is meaningful or predictive according to some external criterion. Some examples of techniques related to classification are discussed in this chapter.

Many analyses of community indicators have the goal of explanation or prediction, which includes discovering why communities differ from one another or uncovering the factors responsible for community change over time. In recent years there has also been considerable interest in determining whether community initiatives have been successful in improving neighborhood or individual outcomes (e.g., impact studies) and in how neighborhoods affect residents (e.g., neighborhood effects studies). Both of these are explored in more detail in this chapter. When explanation is the purpose, an important methodological consideration is how to reduce the chances of making biased causal attributions. Several recent methodological advances related to valid methods for neighborhood studies have explanation and causal attribution as their primary purpose.

## Conceptual Focus and Measurement Focus

The concept of neighborhood belies its layers of complexity and the fact that each layer requires a somewhat different approach to conceptualization and analysis. In fact, these layers might be thought of as nested, such that people are nested within households and housing units, housing units and other physical attributes are nested in neighborhoods, and neighborhoods are nested in cities and regions. The focus for an analysis can also be at one of several conceptual levels, such as *people and place attributes, community structure or process,* or *spatial patterns and dynamics.*

A great deal of neighborhood data analysis focuses on the people, housing, and physical attributes that characterize specific places. The gathering and preparation of these types of data are discussed in chapter 3. Although the tabulation of these data are typically straightforward counts and rates, the myriad of data elements often makes interpretation unwieldy. This chapter reports on some advances in multiattribute indexes and classification methods that can aid in the analysis of neighborhood attribute information.

Another focus for measurement is referred to here as community structure or process. These structures or processes are social constructs such as institutional arrangements, economic or political structure, network relationships, and collective properties of the community. Although data from individuals or organizations may go into these measures, the assumption is that the constructs are emergent properties of the place or group. These "eco-measures" are seldom simple tabulations but require validation as higher-level aggregate concepts and measures. Although we do not explicitly focus on ecometrics in this chapter because it has been well covered elsewhere (Raudenbush and Sampson 1999), we do review recent developments in neighborhood delineation and data sparseness that are relevant to the measurement of the collective properties of places.

Finally, spatial patterns and dynamics are of increasing interest with respect to community indicators. Spatial metrics take seriously the concept of scale and use geographical information system (GIS) tools to quantify important dimensions such as distance, contiguity, density, clustering, and spread. Spatial measures can be used to calibrate access to resources, exposure to social or environmental problems, or the spatial behavior of individuals and organizations. These metrics can be extended over time to assess shifting locations and movement. We report on some advances in spatial analysis for neighborhood data in the next major section.

Given our focus on recent advances and promising analytic techniques, we make no claim to covering the breadth of methodological considerations required for neighborhood data analysis. However, the analytic tools and techniques covered below have useful applications across several of the purposes and conceptual focuses discussed above.

## Review of Methodological Tools and Techniques

The methodological review is organized under the following topics:

- *Specifying neighborhood as a social and geographic unit.* There is debate about how the concept of neighborhood should be operationalized for analysis. This section describes some innovative methods of delineating neighborhoods for varying purposes.
- *Addressing data sparseness.* Because they are often small areas, there may be too few data points within some neighborhoods to make reliable measures. This section describes techniques to improve measurement reliability under these circumstances.
- *Combining indicators for multidimensional metrics and classification.* Neighborhoods are seldom one dimensional, and many analytic problems call for methods of data reduction or classification. This section reviews techniques that have been found useful for multidimensional neighborhood analysis.
- *Assessing the geography of resources and disamenities.* Whether residents can conveniently access resources and avoid exposure to negative social and environmental conditions is not simply a matter of determining what is located within a neighborhood. In this section

we examine examples of how neighborhood access to resources and exposure to disamenities can be calibrated.

- *Evaluating the impact of neighborhood initiatives and action.* Many programs and policies are targeted at improving neighborhoods, but standard program evaluation designs have been difficult to apply to these types of efforts. This section reviews some rigorous methods that have been found useful for evaluations at the neighborhood level.
- *Understanding neighborhood effects on individuals and households.* That neighborhoods matter is a premise of much of the neighborhood work described in this volume. However, there remain numerous unanswered questions about how neighborhoods affect the well-being of populations. This section discusses recent methodological advances in seeking answers to these questions.

## Specifying Neighborhood as a Social and Geographic Unit

Data analyses that aim to describe neighborhoods or determine their impacts must first specify the units for data collection and aggregation. However, what constitutes a neighborhood and how to recognize its boundaries for purposes of analysis are longstanding questions (Downey 2006; Galster 2001). If the unit as measured is not the neighborhood area that is pertinent to the issue under investigation, error or bias can be introduced into the findings. Too often, neighborhood indicators are calculated for administrative districts such as census tracts, Zip Codes, or planning areas without considering the consequences for measurement validity. Yet failure to measure at the correct scale for the issue at hand can lead to erroneous conclusions about neighborhood conditions or effects (Messer 2007). This section describes additional methods of delineating neighborhoods that may be more valid for some purposes than the commonly used administrative units.

Resident-defined neighborhoods are one option to consider. Residents' perceptions of neighborhood boundaries may be particularly useful when there is an interest in how neighborhoods are experienced by people. At the individual level, residents' boundary definitions may be a function of their personal identities, social relationships, or travel patterns. In addition, features of the space such as the demographic composition, built environment, and collective identity may also determine how individuals see their neighborhoods. Researchers have confirmed that residents' views do not map neatly onto census tracts and that there may be disagreement among residents as to neighborhood definitions (Campbell et al. 2009; Coulton et al. 2001; Foster and Hipp 2011). However, community mapping exercises and GIS tools hold promise as a way to link perceptions to geographic boundaries in a useful way.

An illustration of how GIS tools can be used to uncover resident-defined neighborhood units comes from the Making Connections program of the Annie E. Casey Foundation (Coulton, Chan, and Mikelbank 2011). Making Connections, described in more detail in chapter 5, was a community change initiative directed toward strengthening neighborhoods for families and children in low-income communities. Representative samples of households in each community were asked to draw the boundaries of their neighborhoods as they viewed them onto cartographic maps. The residents were grouped according to their self-reported neighborhood names, and their digitized maps were overlaid to find areas of consensus. The blocks marked in a plurality of resident maps were considered to be core parts of the neighborhood for the purposes of community identity. Local stakeholders reviewed the resulting neighborhood units and provided some evidence of face validity based on their understanding of the local context. The resident-defined neighborhoods were then used in specifying the units for the aggregation of survey and census data. This process allowed the calculation of neighborhood indicators for the areal units that residents collectively defined as consistent with their sense of neighborhood identity.

Another approach to neighborhood boundary definition is to use aspects of the built environment that structure social processes and everyday life. T-communities are neighborhood units defined by the network of streets that pedestrians can traverse without crossing main streets (Grannis 2005). The boundaries of T-communities are demarcated using street data from the Census Bureau's TIGER line files so that tertiary street networks fall within the area that is bounded by main streets or other physical barriers. GIS tools are used to identify the blocks that fall within these networks of pedestrian streets. The resulting areal units (T-communities) demonstrate predictive validity with respect to various hypotheses about racial segregation and interaction (Grannis 2005). Moreover, T-communities identified through GIS can be combined with local knowledge to further refine this street-based definition of neighborhood units (Foster and Hipp 2011).

An additional option is to use person-centric neighborhoods (sometimes referred to as ego-centric or sliding neighborhoods) that use GIS tools to specify individualized neighborhood boundaries centered on the location of households. Buffers of varying sizes are drawn around each household's location, and neighborhood indicators are calculated for these buffers (Chaix et al. 2009; Guo and Bhat 2007). The buffer may be specified by distance weights, population size, or other geographic features (Chaix, Merlo, and Chauvin 2005). In this way, a neighborhood measure is created for each household's unique area. There is evidence that the magnitude of contextual effects on some health outcomes is greater when ego-centric or sliding rather than the administratively defined neighborhoods used in statistical models (Chaix, Merlo, Subramanian, et al. 2005).

Finally, sometimes it may be desirable to craft neighborhood units that are demographically homogeneous, are of a designated size and shape, or that do not cross selected barriers or landmarks. Automated zone-design programs can be used to aggregate areas while optimizing such criteria (Cockings and Martin 2005). This method of crafting neighborhood units was investigated in Bristol, England, following an iterative process that imposed various constraints with respect to population and housing characteristics, area size, and geographic considerations (Haynes et al. 2007). The resulting neighborhood units were similar to community areas that were designated by local government officers.

All these alternatives to administratively defined neighborhoods face practical challenges. The availability of data can be a problem. Data at the point level or data for small aggregations such as blocks are fairly easy to allocate into unique boundaries, but data that are available only for larger geographies (e.g., census tracts) may have to be approximately apportioned to the new units. Moreover, the burden of analytic work required to use resident maps, street networks, spatial buffers, and so forth is not trivial. However, when the purpose of a study is to explain how neighborhoods change or affect behavior, such methods hold promise as a way of assuring that the neighborhood is correctly specified for the individuals involved.

Finally, although the above techniques are promising as alternatives to the delineation of neighborhoods by administrative agencies, it is important to further evaluate them based on a deeper understanding of how people interact with and relate to their neighborhoods (Matthews 2011). The relevant neighborhood unit will differ depending on the desired outcome and by characteristics of the individual (Spielman and Yoo 2009). Moreover, greater attention is needed to the variation within neighborhoods on outcomes and how the designation of neighborhood units affects the magnitude of within-neighborhood variation (Merlo et al. 2009).

## Addressing Data Sparseness

Neighborhood indicators may rely on data collected on samples rather than the entire universe of residents, housing units, streets, or other entities. The data from the sampled units are then aggregated to summary measures, such as means, medians, and percentages, of the neighborhood. However, when neighborhood sample sizes are small, sample-based estimates of neighborhood characteristics will have margins of error that may be so large as to be problematic either for descriptive estimates or for use in more complicated analysis.

A simple fix for this problem is to combine multiple neighborhoods or several years of data to gain sufficient sample size and, thereby, more precise estimates. However, these aggregations are typically less useful on a practical level, and they tend to introduce unwelcome heterogeneity because important variations may be obscured when time points are combined. Another option is to leverage spatial correlation to improve the reliability of measures. By using information from surrounding areas, the resulting estimates for each place borrow power from the sample in those areas that are nearby.

When an unreliable estimate is adjusted based on the estimates in nearby areas, the new quantity is often referred to as a *shrinkage estimate.* This term reflects the reality that the adjusted estimate shifts toward the surrounding values based on the relative degree of unreliability in the estimates. In other words, an estimate for a neighborhood that has a high standard error (due to small sample size and/or large variance) moves closer to estimates from other areas that have less error. Neighborhood estimates that have little error do not shrink very much. Because the reliability (or error) of the measures is determined from the data, these are referred to as empirical Bayes estimators (Bingenheimer and Raudenbush 2004). There are a few examples of how neighborhood measures have been made more reliable by using estimated values (i.e., mean and variance) at the city or regional level to produce shrunken estimates for the neighborhoods (Mast 2010; Noble et al. 2010).

When the values of neighborhood indicators are spatially correlated, then taking distance into account may provide a more precise approximation and better address the problem of unreliability. The value of considering distance was demonstrated in a study that compared survey-based neighborhood measures using empirical Bayes estimates that were nonspatial to those that took spatial contiguity into account (Savitz and Raudenbush 2009). Using survey data from the Project on Human Development in Chicago Neighborhoods, the researchers selected only a small subsample of the cases to simulate a situation of sample sparseness. A first-order contiguity matrix was used as the model for spatial dependence. Applying a two-level spatial hierarchical linear model to a measure of collective efficacy, they demonstrated that the neighborhood estimates using spatial dependence have less error and more predictive validity than do unadjusted estimates or nonspatial empirical Bayes estimates. Mujahid et al. (2008) also provide a straightforward illustration of how spatial dependence can be leveraged to improve the reliability of census tract measures based on sample surveys using GeoDa, a spatial data analysis software package.

Too often, neighborhood indicators are based on sparse sample estimates that may be unstable and can be misleading in practice if they are taken at face value. Placing a confidence interval around the estimates is one way to communicate such uncertainty when simple description is the aim. However, for cross-neighborhood comparisons, analysis of trends, or more complicated studies that attempt to uncover relationships among neighborhood measures, the unreliability of the sample-based estimates can attenuate the findings. Shrinkage estimates developed using spatial analysis software can provide more reliable metrics for descriptive and comparative purposes.

## Combining Indicators for Multidimensional Metrics and Classification

Communities vary in innumerable ways, making single indicators of limited usefulness for research and planning. But the presentation of numerous individual indicators at once can be unwieldy and difficult to interpret, especially if communities are not in the same rank order on all of them, or if the indicators fall into several, perhaps overlapping, domains. In addition, combinations of indicators may be stronger or more accurate predictors of community needs or outcomes than single indicators. However, it is not a simple matter to combine indicators so that communities can be compared or classified. Two areas in which there have been important methodological advances are the development of indexes of well-being applicable to small areas and the creation of typologies of places.

There is a long history of efforts to develop indexes that can provide a comprehensive measure of the overall well-being of the population (Hagerty et al. 2001; Hagerty and Land 2007; Land 2000). Most of this work has been at the national level, or has been comparative across nations, and has focused on quality of life across multiple domains. Much has been learned over the years about how to formulate multi-indicator indexes, and many of the principles can be applied to small areas such as neighborhoods. First, each of the indicators that go into the index should be as reliable as possible. Second, the indicators must be standardized or put on a common scale in order to combine them. Third, a decision has to be made about how to weight the indicators relative to one another. They can be weighted equally, but this is still a choice and requires justification. Fourth, it is important to be aware of the degree to which the indicators are correlated with one another and the pattern of the correlations. Depending on the purpose of the index, the indicators do not necessarily need to be correlated, but an index that is multidimensional may need to be decomposed into its subcomponents for additional analysis. Finally, it is desirable that the index be tested against external criteria to determine its validity. These criteria would ideally be both objective (e.g., in predicting trends or events) and subjective (e.g., as compared with the judgment of individuals).

The Child Well-Being Index in England (Bradshaw et al. 2009), which relies on administrative records data collected for all locales in the nation and is estimated for small areas with populations of approximately 1,500, is an example of a carefully developed small-area index. Its basic indicators are rates and proportions. When the rates are deemed unreliable due to small numbers, shrinkage estimates (see above) are used to reduce the margin of error (Noble et al. 2006). All indicators are standardized before being summed into one of several domains. Within domains, varying weights are applied. In some domains, equal weights are applied and in others, the weights are generated through factor analysis. The domains are then summed to form an index that allows all areas to be ranked.

A challenge for any indexes of well-being is to establish the correct weights for combining the individual indicators (Noble et al. 2006). When this is done by factor analysis, the weights come from the overall contribution of each indicator to the common variance among them. If the common variance can be assumed to capture an important latent construct, then these weights make sense because higher weight is given to those indicators most closely related to that construct. Another method of establishing weights empirically is to estimate a statistical model in which the weights for the indictors are derived according to an estimate of each indicator's contribution (e.g., regression coefficient) to some criterion that is valued. Alternatively, it is possible to use public or expert opinion to provide weights for indicators (Hagerty and Land 2007).

Indexes are particularly useful for describing small areas relative to one another. Because they capture a broad construct rather than a single indicator, they are often used to document inequality (Powell et al. 2007) or to target resources to particular areas based on their relatively high scores on an index of need. Applied in this way, indexes are not just research exercises but have important consequences in the real world. Yet the complexity of multidimensional indexes makes them sensitive to the accuracy of the indicators that go into them and the methods used to combine indicators into a single ranking or score. Relatively small changes in methods or assumptions may yield quite different rankings, and mistakes in this regard can produce misleading results. For example, some neighborhood rankings that appear in the popular media combine multiple indicators without sufficient information on the details of how they are produced. It is incumbent on those who use such rankings to demand transparency with respect to the analytic methods underlying them and the evidence for validity of the resulting indexes.

In contrast to multidimensional indexes, neighborhood typologies use multiple indicators to classify places into types, and neighborhoods are differentiated categorically rather than along a continuum. Typologies are formed through multivariate classification schemes, so that the neighborhoods within each type share key attributes and differ on those attributes from neighborhoods of other types. Simply put, typologies identify mutually exclusive groups of entities that are more similar within the groups than between the groups. Unlike multi-indicator indexes that attempt to rank neighborhoods along some continuum, neighborhood typologies can be thought of as categorical.

Typologies are particularly useful for revealing the intersection of social and economic factors that shape the differentiation among neighborhoods. Although neighborhood types are by definition somewhat idealized, they serve to synthesize a multidimensional set of differences into a smaller number of groups based on the observed patterns in the data.

Indeed, since it is often a combination of factors that interacts to advantage or disadvantage particular communities relative to others, typologies are a useful starting point for studying the dynamics of urban structure or the impact of place on the well-being of the population. Such classifications serve many practical purposes as well, such as describing patterns of change over space and time and tailoring public policy decisions to the unique conditions of places that are of different types.

Meaningful classification of neighborhoods is a data- and analytically intensive exercise. The indicators used in the investigation are determined by the purpose of the typology and relevant theories about the processes driving the differences among places. The variables selected for inclusion represent relevant distinctions among places. Though each neighborhood has a unique profile on the set of variables, the goal of the analysis is to find a smaller number of groupings that account for the patterns in the data. With even a few variables, the number of possible profiles becomes impossibly large, so it is necessary to apply analytic methods to uncover meaningful patterns of similarities and differences.

Cluster analysis has been widely used in urban affairs to classify cities, suburbs, and rural areas for a variety of purposes (Mikelbank 2004). This technique has also been applied to the classification of small areas such as communities and neighborhoods. (Cluster analysis to identify types should not be confused with spatial clustering methods, which are discussed in another section of this chapter.) Chow (1998), for example, found that four clusters adequately differentiated among neighborhoods in Cleveland, Ohio, on 10 social problem indicators. Stable neighborhoods had low scores on all social problems. Transition neighborhoods were beginning to show health problems but had low rates of other problems. Distressed areas had economic problems but still remained relatively safe, and extremely distressed neighborhoods had high rates on all social problems.

A typology generated through cluster analysis of Chicago neighborhoods was found to be useful in a study of Chicago's New Communities program (Greenberg et al. 2010). (This initiative is also discussed in chapter 5.) The cluster analysis, which used indicators of neighborhood demographics and economic conditions at the start of the program, revealed five types of neighborhoods. This typology was helpful for interpreting early program results because the clusters were used to identify similar types of neighborhoods for comparison on trends in quality-of-life indicators such as crime and foreclosure rates.

Indexes and typologies are practical tools for simplifying a large amount of information about communities and neighborhoods into a few metrics that can support decisionmaking. Although combining measures results in the loss of some detail and nuance, the intersection of conditions revealed through indexes and typologies often provides a better understanding of a community's assets and challenges. Caution is in order, however, because if measures are not combined with adequate care, the results can be misleading and unstable. Care must be taken to assess the assumptions underlying the statistical methods to see that they are met in practice. Measurement error, highly correlated variables, and outliers are common in community indicators data, and these can skew the results. Moreover, findings may not be stable across time periods or communities.

In an essay accompanying this chapter, Bodini provides an important example of how many of the methods of analysis described above can be rigorously applied to neighborhood indicators. He describes the development of a dynamic neighborhood typology that is longitudinal and multidimensional and illustrates how the typology and related tools can be used to understand and predict neighborhood change.

## Assessing the Geography of Resources and Disamenities

Neighborhoods are by definition places that have a geographic location. Therefore, spatial metrics and analysis have tremendous potential to contribute to neighborhood measures and the analysis of neighborhood conditions and processes. However, a great deal of the analysis of neighborhood indicators and research on neighborhoods has treated them as nonspatial social units, thereby missing the influence of proximity and geographic access that are part of the neighborhood experience (Downey 2006; Spielman and Yoo 2009). Neighborhood indicators that are explicitly spatial can provide more refined information about a place and the advantages and disadvantages of its location. Moreover, spatial analysis can overcome some of the limitations of small-area data, such as arbitrary boundaries, data sparseness, spatial heterogeneity, and spatial dependence. This section reports on several applications of spatial metrics and analysis that have particular relevance to neighborhood indicators work.

Characterizing neighborhoods in terms of their proximity to opportunity or risk is of longstanding interest in the field. One of the reasons place is theorized to matter is that it can confer advantages associated with access to services, jobs, amenities, and so forth. Places can be disadvantaged by the inaccessibility of such resources, and they can also be negatively affected by their propinquity to environmental toxins, noise, disorder, and other hazards. Although it may be relatively simple to quantify the resources or hazards that lie within the boundaries of particular neighborhoods, a more valid approach is to take distance into account in a more continuous fashion. GIS tools have been successfully used to produce distance-based measures that are useful for neighborhood indicators.

In order to craft spatially calibrated measures of access or exposure, it is first necessary to consider the nature of the phenomenon in question and how it is manifested in space. This understanding will guide a number of methodological decisions that must be made. First, the data analyst must decide how distances to resources or hazards are to be calculated. Frequently, linear distances are used. However, if travel time is a concern, it may be more appropriate to calculate distances along roadways or public transportation routes. For some purposes a threshold may be established, such as whether a resource is within a specified distance or beyond a particular boundary. Also, a method of distance weighting must be considered. For example, it is possible to give more weight to resources that are closer to the neighborhood than to those that are farther away by using weights that follow a distance decay function. Second, the spatial granularity of the data will affect the precision of exposure or access indicators. When data on locations are in an aggregated form (e.g., number of resources in each census tract), an assumption will need to be made about where to place the locations within the unit (e.g., at the centroid, randomly distributed throughout the area, and so forth). Aggregated data may also need to be weighted for the size of the aggregation unit so that small units do not have undue influence on various calculations. Several applications are discussed below that illustrate some of these alternative specifications and how they fit the purpose of the analysis.

A study of access to mental health and substance abuse services in the Detroit, Michigan, area illustrates the use of a buffer with unweighted, aggregated count data (Allard, Rosen, and Tolman 2003). The researchers built a geocoded database of all providers that served low-income individuals in the metro area. Based on interviews with experts, they established a 1.5-mile buffer as a definition of adequate service access. For each census tract in the study, they calculated the number of providers within the 1.5-mile buffer and standardized the score by dividing by the mean count for all tracts. The service access scores were not adjusted for competition (i.e., the number of individuals in the buffer eligible for services), nor were providers outside the buffer given any weight in the access score. Among other things, the study found that low-income welfare recipients living in the Detroit suburbs had better access to mental health providers than did those inside the city limits.

A second illustration comes from a study of neighborhood access to entry-level job opportunities in the Cleveland metropolitan area (Bania, Leete, and Coulton 2008). A job access measure was calculated for every census tract in the area by using information on the number of job openings and the distance between every pair of census tracts. Three measures of distance were tested: linear distance, travel time by auto, and travel time by public transit. A decay function was used to weight job openings according to their relative proximity. The job access score for each tract consisted of the distance-weighted number of entry-level job openings divided by the weighted number of residents who were considered competitors for entry-level jobs. The score was normalized by dividing by the mean of all tracts and multiplying by 100. The study found that job access measures differed depending on which measure of distance was used.

A third example reflects recent concerns about food deserts and efforts to develop measures of geographic access to healthy foods. A study in the Portland, Oregon, metropolitan area illustrates some of the alternative specifications that can be considered for such measures (Sparks, Bania, and Leete 2011). The authors geocoded the locations of supermarkets and grocery stores by using addresses collected from several databases. They also classified the stores with respect to the supermarket chain they belonged to. They calculated several spatial metrics for each neighborhood, including the number of stores within a one-kilometer distance and the distance from each neighborhood to the nearest supermarket. Distances were measured in multiple ways, including Euclidean distances between the block group and census tract centroids and distances along street networks. Based on comparisons of these multiple measures, the authors drew some conclusions pertinent to neighborhood indicators work. One conclusion is that distance-based measures of food access are highly correlated regardless of whether Euclidian or street network distances are used and whether block group or census tract levels of aggregation are applied. However, the measures that used a fixed boundary of one kilometer yielded somewhat different patterns than those that used distance measures, suggesting that if fixed boundaries are used in access measurement they should be carefully justified based on substantive considerations. Food access researchers often use one kilometer because it is considered walkable, but this distance may be more or less appropriate depending on public transit routes, land use, and pedestrian streets, which were not taken into account in this study.

GIS tools can also be used to craft measures of proximity to disamenities. Exposure to hazardous wastes was examined from a spatial perspective for neighborhoods in Detroit (Downey 2006). GIS tools were used to calculate exposure to toxic releases for each census tract so that census variables could also be used in the study. Data on the locations and amounts of toxic releases were obtained from the Environmental Protection Agency's Toxic Release Inventory. In order to convert these point-level release data into exposure measures for census tracts, a raster grid was overlaid on the census tract map. For each cell in the grid, the total number of pounds of pollutants emitted within one-quarter kilometer was calculated. The emission score for each tract was calculated by summing the emissions for the cells in the tract and dividing by the number of tracts. This procedure can be seen as standardizing the emissions for tract size and capturing emissions that are adjacent to the tracts themselves.

A study in New York City included a spatial measure of disamenities that were thought to interfere with park use (Weiss et al. 2011). Because one of the possible factors was violent crime, several years of point data on homicides were obtained from public sources. Homicides are relatively rare events, but they have the potential to induce fear among residents at both the site of occurrence and over time and space. To quantify this pattern of influence for the neighborhoods surrounding parks, the researchers used GIS tools to estimate a spatially smoothed kernel density grid by using inverse distance weighting from the homicide points. An average homicide density was calculated for each neighborhood. The same technique was used to quantify the density of other disamenities, such as traffic accidents. Disamenities were found to be more of a limiting factor for park access in neighborhoods with low socioeconomic status than in more affluent areas.

The above examples demonstrate how GIS tools can be used to craft measures of neighborhood access and exposure that are more precise than simple counts of what is present within neighborhood boundaries. However, many questions remain about the most appropriate buffer sizes and distance decay functions for particular phenomena. The field will benefit from continued research that provides empirical evidence to guide the proper specification of the correct geographic range and scale for the various processes of interest.

## Evaluating the Impact of Neighborhood Initiatives and Action

In many instances questions arise about the degree to which a policy, program, or practice has affected neighborhoods. Such questions may be in conjunction with an evaluation of a government- or foundation-sponsored initiative that is aimed at community improvement, or they may be in response to concerns raised about whether a policy or program is having positive or negative effects at the neighborhood level. These impact questions cannot be answered definitively by simply looking at trends in selected neighborhood indicators. Instead, care must be taken to craft an evaluation design that has a plausible counterfactual; that is, the design must include a way of comparing what would have happened if the initiative, program, or policy under investigation had not occurred. Although the field of program evaluation is well developed with respect to research designs that are valid for this purpose (Rossi, Lipsey, and Freeman 2004), these techniques have proved difficult to apply in practice to questions of community or neighborhood change (Hollister and Hill 1995; Rossi 1999). Most obviously, this difficulty occurs because the tried-and-true methods of randomized controlled trials, which have become the gold standard for human studies, have been difficult to apply when the subjects are neighborhoods (or households nested within neighborhoods) instead of unrelated individuals. In this section we lay out some of the challenges that face evaluators in answering impact questions with respect to neighborhoods and report on several examples of approaches that have overcome some of these difficulties.

Because most program evaluations focus on individuals, it is first useful to discuss some of the unique aspects of place-based program and policy evaluation. An initial question is to consider whether the policy or program is aimed at attributes of the place or at the behaviors or characteristics of the people within the place. Although this sounds like a simple distinction, it is often murky in practice. For example, community development initiatives are often directed toward improving housing conditions or increasing economic activity in the neighborhood. At first glance, these would seem to be attributes of the neighborhood. However, a deeper look suggests that such outcomes are difficult to separate from attributes and behaviors of residents or investors who will be called on to make the housing and business decisions that are necessary to make the programs successful. Communities may also sponsor social programs aimed directly at individuals in the neighborhood as part of a comprehensive approach to community improvement. Individuals who participate in or qualify for the programs may be of interest in this case, but the fact that the program is targeted within a place raises evaluation questions and issues that are not of concern in the typical program evaluation focused on individuals who are assumed to be independent of one another. In fact, part of the evaluation may be to determine whether the participants and the larger community benefit from the fact that a threshold proportion of residents is now participating in a particular program. Programs implemented in this way are sometimes referred to as saturation models.

These complexities call for program theory to guide the evaluation that takes into account direct effects on the participating or targeted entities and also spillover effects on surrounding persons, residential properties, or businesses. Similarly, the evaluation design cannot rely on standard assumptions that prevail in individually focused randomized trials, such as the assumption that the units, whether individuals or neighborhoods, are independent from one another or that they are exchangeable (Merlo et al. 2009; Oakes 2004). Among other things, researchers need to be on the lookout for heterogeneous treatment effects and correlated errors due to spatial proximity or social interaction patterns, both within and between neighborhoods.

A number of practical considerations challenge the evaluator of neighborhood interventions. Enrolling sufficient numbers of neighborhood units to achieve adequate statistical power typically exceeds the costs of enrolling a similar number of individuals. Moreover, although the principles and methods of informed consent are well developed for individuals who agree to be in randomized trials, the experience with enrolling neighborhoods in experiments is quite limited. Indeed, given the many persons and organizations that are stakeholders in the typical neighborhood, it is not surprising that consensus about participation in research is difficult to obtain.

As discussed in chapter 5, a high rate of residential mobility can also complicate efforts to demonstrate the impact of a program or policy at the neighborhood level (Coulton, Theodos, and Turner 2012). Households move frequently even under normal circumstances, and persons participating in a program may have a different probability of leaving the neighborhood and reasons for moving than those who do not participate. Residential mobility may also limit the length of households' exposures to place-based interventions, or the move itself may exert an influence on outcomes. Research designed to evaluate neighborhood initiatives faces the challenge of correctly specifying how mobility may affect impact, including the effect on those who leave the neighborhood. To have a valid counterfactual, it may be necessary to track out-movers, as well as those who stay in place or move in, after a program or policy has gotten underway.

Despite these many challenges, quite a few studies have examined the neighborhood-level effect of policies or programs targeted at individuals, properties, or other entities. Several of these are described below, with a particular focus on the elements of the research design that were used to craft a counterfactual for the evaluation.

### *Cluster Randomized Trials*

We begin with a strong counterfactual design, one in which neighborhoods are randomly assigned either to an experimental group that receives an intervention, program, or policy or to a control group that does not. When the focus of the research is on the impact on individuals (or other units) within neighborhoods, these designs are known as cluster randomized trials. There is a growing interest in such methods as tools for evaluating program impact within the context of place (Boruch 2005; Cook 2005). A number of cluster randomized trials have been conducted in which small areas have been randomly assigned to receive various types of interventions, such as health promotion, targeted policing, or coordinated social services (Boruch et al. 2004).

These designs are appropriate when an intervention is directed at places and the entities (e.g., persons, houses, events) located there. For example, hotspot-focused policing has been shown in cluster randomized trials to reduce crime incidents in the street segments or address points within the hotspots (Weisburd 2005). Cluster randomized trials are also appropriate for evaluating the impact of interventions directed at individuals when theory suggests there is something about the community context that is relevant to the outcomes (Bloom 2005). For example, coordinated community services may be more effective when there has been widespread place-based involvement in the planning process or when neighbors or members of the same organization are collectively involved in the programs. The benefits may spread among participants, or broad-based community engagement may contribute to program quality. Given such assumptions, randomly assigning unrelated individuals to receive services would fail to produce the types of engagement and participation called for by the theory. Organizers of such place-based interventions would undoubtedly anticipate impact on individuals (e.g., increased knowledge or skills), but they would also be interested in showing effects on the community as a whole (e.g., increased collective efficacy).

Given the multilevel structure of cluster randomized trials, various important statistical considerations must be taken into account when planning these studies. One of the most crucial is the question of the statistical power to detect impact at the cluster and individual levels. The power is dependent on several factors. The number of clusters and the number of cases per cluster are key elements of statistical power. Covariates can be added at either level to reduce variance due to pre-existing differences, which also contributes to power calculations. Finally, a consideration that is particular to cluster randomized designs is the role played by the degree of within- and between-cluster heterogeneity (Raudenbush 1997). Although a detailed discussion of statistical and design principles is beyond the scope of this chapter, the W.T. Grant Foundation has undertaken to provide practical tools for researchers interested in implementing cluster randomized trials (see http://www.wtgrantfoundation.org/resources/research-tools).

The evaluation of the Jobs-Plus employment program for public housing residents is an example of a randomized trial applied to evaluation of a neighborhood-level intervention (Bloom and Riccio 2005).The goal of Jobs-Plus was to demonstrate that a place-based and comprehensive employment-focused intervention could raise employment rates among public housing residents. The initiative rested on the premise that focusing financial incentives, employment programs, and resident engagement in a place would be an effective way to address the employment problems of public housing residents. Public housing developments in five cities were randomly assigned to be in an experimental or control group. Jobs-Plus used a comparative interrupted time series design to create a strong counterfactual (Bloom 2005). The researchers were able to construct a multiyear baseline trend and a postintervention trend on employment rates by using data on all adults living in both the experimental and control sites. An additional feature of the Jobs-Plus evaluation was that residents who were exposed to the intervention were tracked even if they left the public housing development. This design allowed the estimation of the causal impact of the place-based Jobs-Plus model on individuals regardless of whether they stayed in public housing the entire time or moved. In addition, it allowed an estimate of the interventions' impact on employment levels in the public housing sites over time.

### *Matched Neighborhood Designs*

Matching is a technique that has been used to craft comparison groups of neighborhoods when randomization is not feasible. This method typically begins with a set of neighborhoods that is involved in a program or policy and then searches for comparison neighborhoods that are as similar as possible to the participating neighborhoods on relevant variables. A limitation of matched designs, as compared to randomly assigned control groups, is that there may be unobserved factors that contribute to the selection of neighborhoods into a program or policy. Randomization equalizes the treatment and control groups on measured and unmeasured factors, but matching is only able to control for variables that are measured. Moreover, perfect matches can seldom be achieved, so the precision of the matching approach and related aspects of the statistical analysis are other methodological issues that must be considered.

Propensity score matching is a powerful tool because it enables neighborhoods to be matched on a large number of covariates in combination. The details of estimating propensity scores, choosing the closest matches, and evaluating the success of the matching are beyond the scope of this article (see Guo and Fraser 2010; Oakes and Johnson 2006). However, a useful example comes from an effort to evaluate the impact of the Empowerment Zone (EZ) program in six cities (Rich and Stoker 2014). EZ was a federal initiative to increase economic opportunity in selected urban areas by supporting comprehensive community and economic development, community engagement, and institutional capacity building. Cities chose clearly defined target areas made up of eligible census tracts to implement their programs. Because only a portion of the eligible tracts in each city were able to be included in the EZ, the researchers chose an evaluation design that compared EZ tracts with matched non-EZ tracts on selected economic outcomes. The application of propensity score matching was successful in eliminating most statistically significant pretreatment differences between treatment and comparison census tracts. Moreover, the study revealed heterogeneity in treatment effects within and between sites that had not been fully appreciated based on previous studies that relied on pooled samples and econometric modeling rather than matching as a way to address selection bias (Rich and Stoker 2014).

### *Regression Discontinuity Design*

The regression discontinuity design can be applied to the evaluation of the impact of policies on neighborhoods under some specific circumstances. This design is useful when evaluating a program that has a clear cutoff for neighborhood eligibility, so that neighborhoods on one side of the threshold receive the treatment and the rest do not. If the cut point cannot be manipulated, selection bias is ruled out by the design. A discontinuity in outcomes for neighborhoods on either side of the cut point is taken as evidence of program impact. Deng and Freeman (2011) consider the usefulness of this design for evaluating place-based programs by taking up the example of low-income housing tax credits. Their example illustrates the practical problems of implementing the design in a program that has multiple, overlapping criteria for qualifying neighborhoods. They conclude that the regression discontinuity design is promising for evaluating place-based programs, especially if the need for a precise eligibility cutoff is built into the program implementation.

### *Interrupted Time Series Designs*

Another method of evaluating the impact of policies or programs on neighborhoods is to compare levels and trends on some outcome for intervention and comparison areas before and after a program is implemented. Simple before and after comparisons are not very convincing because areas targeted for a program may have been on different trajectories before the program even started than those not selected. Outside events, contemporaneous with the start of a program, may also have affected outcomes, obscuring the actual effect of the program. With respect to the evaluation of community development programs, Galster and his colleagues (2004) suggest that interrupted time series designs can be strengthened if repeated measures can be made on selected outcomes in intervention and comparison neighborhoods before and after the program begins. Deviations from pretreatment trends in the intervention neighborhoods as compared to the comparison neighborhoods (i.e., differences in differences) can provide convincing evidence of program impact. Having a relatively long and stable baseline is an advantage in terms of the statistical power of these comparisons. And as with any method, the analysis is strengthened by including appropriate covariates that affect the outcomes. Known as the adjusted interrupted time series method, this type of design has been applied in neighborhood studies mainly to address the impact of various policies and practices on housing values or crime because of the long time series that can be developed from police and property records.

### *Localizing Outcomes in Space*

When researchers apply experimental designs that have come out of human subjects research to neighborhoods, they typically treat the neighborhood units as nongeographic entities. As such, neither the location of the neighborhoods relative to other places nor the geographic subareas within them are taken into account. However, spatial concepts can be used to craft more refined counterfactuals, especially when there is reason to suspect that the impact of an intervention has some type of spatial parameters. Increasingly, researchers use GIS tools and spatial locations to calibrate where exposures and effects are occurring. Combined with interrupted time series or comparison group methods, spatial analysis can bolster the validity of the designs.

For example, a study of the impact of supportive housing developments on crime rates in Denver, Colorado, neighborhoods drew buffers of varying sizes around supported housing units and compared crime rate trends in these zones with trends farther away (Galster et al. 2002). Similarly, a study of the effect of dispersed public housing on home values in Denver also examined sales prices around public housing units at various geographic scales and compared them to areas without public housing within similar distances (Santiago, Galster, and Tatian 2001). In New York City, researchers examined the relative effect of housing rehabilitation carried out by commercial and nonprofit developers on sales prices. They also examined whether the differences between these two types of development depended on the size of the buffer used in the analysis (Ellen and Voicu 2006).

There is growing awareness that scale matters in neighborhoods and that the effect of particular policies or practices may vary both within and between so-called neighborhoods. A study in Seattle of the impact of subsidized housing on nearby property values found that the effects of the program were quite different across subareas, and these differences were related to a number of geographically specific local conditions (Koschinsky 2009). Moreover, the location of subareas relative to areas that were more advantaged or disadvantaged was also important. Statistical models that estimated average effects of subsidized units were found to obscure this spatial heterogeneity of program impact.

In an essay accompanying this chapter, Koschinsky argues that evaluation research on neighborhood interventions needs to go much further to incorporate spatial metrics and dynamics into the research design and analysis. She provides illustrations showing how geographic concepts and spatial analysis methods incorporated into place-based evaluations can potentially yield more robust and nuanced estimates of program and policy impact than evaluations that do not apply spatial analysis techniques.

## Understanding Neighborhood Effects on Individuals and Households

A great deal of neighborhood indicators work is motivated by the belief that neighborhoods matter for the people who live in them, yet this belief remains a remarkably contentious proposition among researchers and policymakers. Several important articles have reviewed the evidence for neighborhood effects and identified various theoretical and methodological limitations (Ellen and Turner 1997; Leventhal and Brooks-Gunn 2000; Oakes 2004; Sampson, Morenoff, and Gannon-Rowley 2002). Although there is no doubt that markers of human health and well-being differ by neighborhood, it has proved difficult to establish the causal effects of neighborhoods due to several challenges. One major problem is neighborhood selection, which makes it difficult to separate the effects of neighborhood factors from individual attributes and choices that determine where people live. In addition, people change neighborhoods and neighborhoods change over time, but studies are seldom able to specify both contemporaneous and cumulative neighborhood influences. Reverse causality is also a possibility that cannot always be ruled out when studies have limited information about the timing of individual outcomes in relation to neighborhood change. Although neighborhood indicators could tap any and all of these processes, naivety about causal process and direction can lead to incorrect interpretations and make it difficult to predict what types of policies would be able to help both people and places under the circumstances.

Although the challenges of establishing causality in neighborhood effects research are widely appreciated, the number of practical solutions to date is limited. One important innovation that has garnered a lot of attention is the use of housing vouchers allocated by lottery to induce residential mobility to better neighborhoods. This exogenous influence, like random assignment in experiments, provides leverage on the problems of selection and reverse causality. The Moving to Opportunity experiment was designed with these methodological issues in mind (Goering and Feins 2003). Residents of public housing in five cities were offered an opportunity to be randomized into (1) a treatment group that received a housing voucher and counseling to help residents move to a neighborhood with less than 10 percent poverty, (2) a group that received a housing choice voucher that could be used anywhere (Section 8 control), or (3) a control group that did not receive a voucher. Numerous studies have been published from follow-up data collected over many years following randomization to assess the impact of the treatment (i.e., a voucher to move to a low-poverty neighborhood) on outcome measures for children and adults in these households. The findings regarding the impact on outcomes have been mixed. Despite the treatment group living in higher-quality housing and neighborhoods with lower poverty and crime, no significant differences between treatment and control groups were found for household employment levels and income (Sanbonmatsu et al. 2012) or children's educational success (Sanbonmatsu et al. 2006). With respect to heath, adults in the treatment groups compared favorably with controls on a number of health measures (Sanbonmatsu et al. 2012). Health benefits for children were less uniform, with the most consistent positive results being on behavioral health outcomes for female, but not male, adolescents (Kling, Ludwig, and Katz 2005). Researchers are continuing to mine the Moving to Opportunity data to tease out the nuances of the effects and the mechanisms responsible for them.

Although the randomized mobility experiments have the important advantage of removing selection bias from impact estimates, they have several limitations (Sampson 2008). An important limitation for those interested in improving neighborhoods is that the households eligible for these experiments had to live in public housing prior to randomization, but public housing residents represent only a small portion of the population living in disadvantaged circumstances. In addition, even though the experimental group moved to lower-poverty neighborhoods, they tended to stay within the racially segregated areas of their cities rather than being exposed to suburban areas or ones with predominately white populations. Finally, families assigned to the voucher group did not necessarily move, and many families in the control group did not stay in their initial unit. Moreover, movers often moved again after their initial assignment, limiting their exposure to the lower-poverty condition.

Given the practical limitations on random assignment of households to neighborhoods, the search continues for methodological improvements that have the potential to address the problem of selection bias while achieving broader population representation and accurately reflecting the processes of residential mobility. Sharkey (2012) demonstrates how longitudinal survey data can be used to model the effect of living in disadvantaged neighborhoods by exploiting the change in concentrated disadvantage that occurs over time in neighborhoods. He defines the equivalent of the experimental treatment as living in a neighborhood that decreased in concentrated disadvantage over a 10-year period. The control condition is defined as living in a census tract that stayed the same or went up in concentrated disadvantage. To control for selection bias into neighborhoods, the experimental and control groups are rigorously matched at the beginning of the 10-year window on neighborhood disadvantage in the preceding 10-year period. The outcomes for the individuals are observed during the 10 years following the treatment. Among other things, this study demonstrates that it is possible in a national sample to observe sizable neighborhood changes (e.g., those in the treatment group saw their neighborhoods fall by more than one standard deviation in concentrated disadvantage) and to model the effect of neighborhood improvement on individual outcomes after neighborhood selection bias has been taken into account.

As illustrated above, considerable scientific research effort has been directed at trying innovative techniques to estimate the impact of neighborhood disadvantage on individuals, net of their residential choices and mobility constraints. However, residential mobility, neighborhood selection, neighborhood effects, and neighborhood change are interrelated processes that are of considerable importance because they not only affect individuals and households, but at the aggregate level they are responsible for the social transformation of neighborhoods and cities. In their essay that follows this chapter, Galster and Hedman argue that rather than narrowly focusing on making neighborhood selection ignorable, which tends to require designs that limit the ability to generalize study results, researchers need to investigate the related processes of residential selection and neighborhood effects. Toward this end, they provide a synthetic model that lays the groundwork for future research that examines residential mobility and neighborhood effects holistically.

## Conclusion

The field has made a great deal of progress in the analysis of neighborhood data and the investigation of questions that are useful in community change. The use of the analytic methods discussed in this chapter further several goals, such as building knowledge about neighborhoods, evaluating programs and practices, providing information to guide planning and action, and advocating for social justice. Some methodological challenges have received a great deal of attention while others are just beginning to be addressed, but work continues to evolve on all the applications covered in this chapter. Looking broadly across the data analysis enterprise, several general conclusions can be drawn.

First, the breadth and sophistication of neighborhood measurement have been advancing at a steady pace, but the fact that the work is scattered across disciplines and substantive areas presents barriers to its application. Relatively few neighborhood concepts have widespread use, uniform definitions, and standardized measures; collective efficacy and concentrated disadvantage (Sampson, Raudenbush, and Earls 1997) are notable exceptions. Many other neighborhood analyses rely on measures that have limited use in a few studies or data sources or have not been fully tested. Furthermore, although many researchers recognize the problem of representing neighborhoods as areal units, the majority of applications still use administrative boundaries to aggregate data for neighborhood measurement rather than alternatives that may be more valid for the particular study question. As yet, there does not seem to be a literature that catalogs and provides critical assessment of neighborhood measurement methods across the board.

Second, there is significant potential to advance the usefulness of neighborhood data through more fine-grained and dynamic spatial analysis, but data sources limit analysts' ability to do this. When source data are at the point level, researchers have the flexibility to examine spatial processes and interrelated changes at various distance specifications within and between neighborhoods. Unfortunately, there is unevenness in terms of data availability at the point level, with most examples of point-identified data coming from housing or law enforcement sources. Social and demographic data at the point level are hard to come by because they are usually based on samples from surveys, such as the American Community Survey. The advent of geographically enabled mobile devices holds great promise for generating spatially granular data on social relationships and activities that can be examined for their configurations in time and space to inform the understanding of neighborhood dynamics.

Third, it is time to have rising expectations for building the evidence base for practices and policies that aim to improve neighborhoods and benefit residents. This chapter gives examples of the clever application of research designs that provide leverage on the problem of the counterfactual and begin to provide greater confidence about the evidence of impact. In addition, analyses that explore spatial variation in outcomes are promising for yielding more nuanced understanding of program effects. It is true there has been some discouragement about the ability to evaluate comprehensive approaches to community change in total because they are community driven and have numerous components. However, the ability to rigorously evaluate limited and controlled innovations within existing community programs is feasible and probably can be done at reasonable cost if the outcomes can be evaluated using available data. Indeed, the idea of low-cost randomized trials related to social spending is gaining traction at the national level, and we can anticipate further development of the methodologies to support such work in neighborhoods.

Fourth, the time is ripe to invest in analytic methods and data sources that can contribute to a better understanding of the dynamic processes that shape neighborhoods and the experiences of residents as they traverse time and space. As pointed out in this volume by Galster and Hedman in their discussion of residential selection and neighborhood effects, research has tended to focus on isolating these processes rather than modeling them holistically. However, data sources that are up to the task of dynamic modeling are rapidly evolving, especially as more longitudinal datasets are being created from administrative records or mobile devices that capture locations at frequent or continuous time intervals. Recent developments in computational social sciences, including data mining, agent-based modeling, and simulation, are also promising and need to be evaluated to determine how they can be applied in new ways to examine questions of neighborhood impact and change (O'Sullivan

2009; Torrens 2010). These developments will be important to enhancing the field's ability to fully evaluate programs and policies that target neighborhood conditions, residential choices, and behavioral outcomes within metropolitan areas.

Finally, the current institutional infrastructure and communication channels do not maximize the synergy and practical impact of emerging methodological developments. The work on new analysis techniques and tools has come out of universities, think tanks, governments, and local intermediaries working with neighborhood data; the effort involves individuals from many scientific disciplines, policy arenas, and practice areas. In some ways this breadth and diversity of perspectives have contributed to an openness and vitality in the enterprise. At the same time, however, it is tough for anyone to keep track of the many developments that might be useful on the ground. Moreover, there are few venues in which the individuals involved broadly in the applications of these methods meet and exchange ideas. Importantly, the conversations tend to occur within substantive areas, scientific disciplines, or practice communities. Yet there are many potential points of connection across the spheres that are not being made. Moving forward on the opportunities and challenges identified in this chapter will require a new network that fosters methodological development in the community information field; such changes are detailed in our recommendations in chapter 8.

## REFERENCES

Allard, S.W., D. Rosen, and R.M. Tolman. 2003. "Access to Mental Health and Substance Abuse Services among Women Receiving Welfare in Detroit." *Urban Affairs Review* 38 (6): 787–807.

Bania, N., l. Leete, and C.J. Coulton. 2008. "Job Access, Employment and Earnings: Outcomes for Welfare Leavers in a US Urban Labour Market." *Urban Studies* 45 (11): 2179–2202.

Bingenheimer, J.B., and S.W. Raudenbush. 2004. "Statistical and Substantive Inferences in Public Health: Issues in the Application of Multilevel Models." *Annual Review of Public Health* 25 (1): 53–77.

Bloom, H.S. 2005. "Randomizing Groups to Evaluate Place-Based Programs." In *Learning More from Social Experiments: Evolving Analytic Approaches,* edited by H.S. Bloom (115–72). New York: Russell Sage Foundation.

Bloom, H.S., and J.A. Riccio. 2005. "Using Place-Based Random Assignment and Comparative Interrupted Time-Series Analysis to Evaluate the Jobs-Plus Employment Program for Public Housing Residents." *Annals of the American Academy of Political and Social Science* 599 (1): 19–51.

Boruch, R. 2005. "Preface: Better Evaluation for Evidence-Based Policy: Place Randomized Trials in Education, Criminology, Welfare, and Health." *Annals of the American Academy of Political and Social Science* 599 (1): 6–18.

Boruch, R., H. May, H. Turner, J. Lavenberg, A. Petrosino, D. De Moya, and E. Foley. 2004. "Estimating the Effects of Interventions That Are Deployed in Many Places: Place-Randomized Trials." *American Behavioral Scientist* 47 (5): 608–33.

Bradshaw, J., M. Noble, K. Bloor, M. Huby, D. McLennan, D. Rhodes, and K. Wilkinson. 2009. "A Child Well-Being Index at Small Area Level in England." *Child Indicators Research* 2 (2): 201–19.

Campbell, E., J.R. Henly, D.S. Elliott, and K. Irwin. 2009. "Subjective Constructions of Neighborhood Boundaries: Lessons from a Qualitative Study of Four Neighborhoods." *Journal of Urban Affairs* 31 (4): 461–90.

Chaix, B., J. Merlo, and P. Chauvin. 2005. "Comparison of a Spatial Approach with the Multilevel Approach for Investigating Place Effects on Health: The Example of Healthcare Utilization in France." *Journal of Epidemiology and Community Health* 59 (6): 517–26.

Chaix, B., J. Merlo, D. Evans, C. Leal, and S. Havard. 2009. "Neighbourhoods in Eco-Epidemiologic Research: Delimiting Personal Exposure Areas. A Response to Riva, Gauvin, Apparicio and Brodeur." *Social Science and Medicine* 69 (9): 1306–10.

Chaix, B., J. Merlo, S.V. Subramanian, J. Lynch, and P. Chauvin. 2005. "Comparison of a Spatial Perspective with the Multilevel Analytical Approach in Neighborhood Studies: The Case of Mental and Behavioral Disorders Due to Psychoactive Substance Use in Malmo, Sweden, 2001." *American Journal of Epidemiology* 162 (2): 171–82.

Chow, J. 1998. "Differentiating Urban Neighborhoods: A Multivariate Structural Model Analysis." *Social Work Research* 22 (3): 131–42.

Cockings, S., and D. Martin. 2005. "Zone Design for Environment and Health Studies Using Pre-Aggregated Data." *Social Science and Medicine (1982)* 60 (12): 2729–42.

Cook, T. 2005. "Emergent Principles for the Design, Implementation, and Analysis of Cluster-Based Experiments in Social Science." *Annals of the American Academy of Political and Social Science* 599 (1): 176–98.

Coulton, C.J., T. Chan, and K. Mikelbank. 2011. "Finding Place in Community Change Initiatives: Using GIS to Uncover Resident Perceptions of Their Neighborhoods." *Journal of Community Practice* 19 (1): 10–28.

Coulton, C.J., J. Korbin, T. Chan, and M. Su. 2001. "Mapping Residents' Perceptions of Neighborhood Boundaries: A Methodological Note." *American Journal of Community Psychology* 29 (2): 371–83.

Coulton, C.J., B. Theodos, and M.A. Turner. 2012. "Residential Mobility and Neighborhood Change: Real Neighborhoods under the Microscope." *Cityscape: A Journal of Policy Development and Research* 14 (3): 55–90.

Deng, L., and L. Freeman. 2011. "Planning for Evaluation: Using Regression Discontinuity to Evaluate Targeted Place-Based Programs." *Journal of Planning Education and Research* 31 (3): 308–18.

Downey, L. 2006. "Using Geographic Information Systems to Reconceptualize Spatial Relationships and Ecological Context." *American Journal of Sociology* 112 (2): 567–612.

Ellen, I.G., and M.A. Turner. 1997. "Does Neighborhood Matter? Assessing Recent Evidence." *Housing Policy Debate* 8 (4): 833–66.

Ellen, I.G., and I. Voicu. 2006. "Nonprofit Housing and Neighborhood Spillovers." *Journal of Policy Analysis and Management* 25 (1): 31–52.

Foster, K.A., and J.A. Hipp. 2011. "Defining Neighborhood Boundaries for Social Measurement: Advancing Social Work Research." *Social Work Research* 35 (1): 25–35.

Galster, G. 2001. "On the Nature of Neighbourhood." *Urban Studies* 38 (12): 2111–24.

Galster, G., K. Pettit, A. Santiago, and P. Tatian. 2002. "The Impact of Supportive Housing on Neighborhood Crime Rates." *Journal of Urban Affairs* 24 (3): 289–315.

Galster, G., K. Temkin, C. Walker, and N. Sawyer. 2004. "Measuring the Impacts of Community Development Initiatives." *Evaluation Review* 28 (6): 502–38.

Goering, J.M., and J.D. Feins. 2003. *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment.* Washington, DC: Urban Institute.

Grannis, R. 2005. "T-Communities: Pedestrian Street Networks and Residential Segregation in Chicago, Los Angeles, and New York." *City and Community* 4 (3): 295–321.

Greenberg, D., N. Verma, K. Dillman, and R. Chaskin. 2010. *Creating a Platform for Sustained Neighborhood Improvement: Interim Findings from Chicago's New Communities Program.* New York: MDRC.

Guo, J.Y., and C.R. Bhat. 2007. "Operationalizing the Concept of Neighborhood: Application to Residential Location Choice Analysis." *Journal of Transport Geography* 15 (1): 31–45.

Guo, S., and M.K. Fraser. 2010. *Propensity Score Analysis: Statistical Methods and Applications.* Thousand Oaks, CA: Sage Publications.

Hagerty, M.R., R.A. Cummins, A.L. Ferriss, K. Land, A.C. Michalos, M. Peterson, and J. Vogel. 2001. "Quality of Life Indexes for National Policy: Review and Agenda for Research." *Social Indicators Research* 55 (1): 1–96.

Hagerty, M.R., and K.C. Land. 2007. "Constructing Summary Indices of Quality of Life." *Sociological Methods and Research* 35 (4): 455–96.

Haynes, R., K. Daras, R. Reading, and A. Jones. 2007. "Modifiable Neighbourhood Units, Zone Design and Residents' Perceptions." *Health and Place* 13 (4): 812–25.

Hollister, R.G., and J. Hill. 1995. "Problems in the Evaluation of Community-Wide Initiatives." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Context,* edited by J.P. Connell, A.C. Kubish, L.B. Schorr, and C.H. Weiss (127–72) Washington, DC: Aspen Institute.

Kling, J.R., J. Ludwig, and L.F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 120 (1): 87–130.

Koschinsky, J. 2009. "Spatial Heterogeneity in Spillover Effects of Assisted and Unassisted Rental Housing." *Journal of Urban Affairs* 31 (3): 319–47.

Land, K. 2000. "Social Indicators." In *Encyclopedia of Sociology,* edited by E.F. Borgatta and R.V. Montgomery (2682–90). New York: Macmillan.

Leventhal, T., and J. Brooks–Gunn. 2000. "The Neighborhoods They Live In: The Effects of Neighborhood Residence on Child and Adolescent Outcomes." *Psychological Bulletin* 126 (2): 309–37.

Mast, B.D. 2010. "Measuring Neighborhood Quality with Survey Data: A Bayesian Approach." *Journal of Policy Development and Research* 12 (3): 113–42.

Matthews, S.A. 2011. "Spatial Polygamy and the Heterogeneity of Place: Studying People and Place via Egocentric Methods." In *Communities, Neighborhood and Health,* edited by L. Burton (35–55). New York: Springer.

Merlo, J., H. Ohlsson, K.F. Lynch, B. Chaix, and S.V. Subramanian. 2009. "Individual and Collective Bodies: Using Measures of Variance and Association in Contextual Epidemiology." *Journal of Epidemiology and Community Health* 63 (12): 1043–48.

Messer, L.C. 2007. "Invited Commentary: Beyond the Metrics for Measuring Neighborhood Effects." *American Journal of Epidemiology* 165 (8): 868–71.

Mikelbank, B.A. 2004. "A Typology of U.S. Suburban Places." *Housing Policy Debate* 15 (4): 935–64.

Mujahid, M.S., A.V. Diez Roux, J.D. Morenoff, T.E. Raghunathan, R.S. Cooper, H. Ni, and S. Shea. 2008. "Neighborhood Characteristics and Hypertension." *Epidemiology* 19 (4): 590–98.

Noble, M., H. Barnes, G. Wright, and B. Roberts. 2010. "Small Area Indices of Multiple Deprivation in South Africa." *Social Indicators Research* 95 (2): 281–97.

Noble, M., G. Wright, G. Smith, and C. Dibben. 2006. "Measuring Multiple Deprivation at the Small-Area Level." *Environment and Planning A* 38 (1): 169–85.

Oakes, J.M. 2004. "The (Mis)estimation of Neighborhood Effects: Causal Inference for a Practicable Social Epidemiology." *Social Science and Medicine* 58 (10): 1929–52.

Oakes, J.M., and P.J. Johnson. 2006. "Propensity Score Matching for Social Epidemiology." In *Methods in Social Epidemiology,* edited by J.M. Oakes and J.S. Kaufman (370–92). San Francisco: John Wiley and Sons.

O'Sullivan, D. 2009. "Changing Neighborhoods—Neighborhoods Changing: A Framework for Spatially Explicit Agent-Based Models of Social Systems." *Sociological Methods and Research* 37 (4): 498–530.

Powell, J.A., J. Reece, C. Rogers, and S. Gambhir. 2007. *Communities of Opportunity: A Framework for a More Equitable and Sustainable Future for All.* Columbus, OH: Kirwan Institute for the Study of Race and Ethnicity.

Raudenbush, S.W. 1997. "Statistical Analysis and Optimal Design for Cluster Randomized Trials." *Psychological Methods* 2 (2): 173–85.

Raudenbush, S.W., and R.J. Sampson. 1999. "Ecometrics: Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology* 29 (1): 1–41.

Rich, M.J., and R.P. Stoker. 2014. *Collaborative Governance for Urban Revitalization: Lessons from the Empowerment Zone.* Ithaca, NY: Cornell University Press.

Rossi, P.H. 1999. "Evaluating Community Development Programs: Problems and Prospects." In *Urban Problems and Community Development,* edited by R.F. Furgeson and W.T. Dickens (521–67). Washington, DC: Brookings Institution.

Rossi, P., M.W. Lipsey, and H.E. Freeman. 2004. *Evaluation: A Systematic Approach,* 7th ed. Thousand Oaks, CA: Sage.

Sampson, R.J. 2008. "Moving to Inequality: Neighborhood Effects and Experiments Meet Social Structure." *American Journal of Sociology* 114 (1): 189–231.

Sampson, R.J., J.D. Morenoff, and T. Gannon-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28 (1): 443–78.

Sampson, R.J., S.W. Raudenbush, and F. Earls. 1997. "Neighborhood and Violent Crime: A Multi-Level Study of Collective Efficacy." *Proceedings of the National Academy of Sciences* 277: 918–24.

Sanbonmatsu, L., J.R. Kling, G.J. Duncan, and J. Brooks-Gunn. 2006. "Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment." *Journal of Human Resources* 41 (4): 649–91.

Sanbonmatsu, L., N.A. Potter, E. Adam, G.J. Duncan, L.F. Katz, R.C. Kessler, … T.W. McDade. 2012. "The Long-Term Effects of Moving to Opportunity on Adult Health and Economic Self-Sufficiency." *Journal of Policy Development and Research* 14 (2): 109–36.

Santiago, A., G. Galster, and P. Tatian. 2001. "Assessing the Property Value Impacts of the Dispersed Housing Subsidy Program in Denver." *Journal of Policy Analysis and Management* 20 (1): 65–88.

Savitz, N.V., and S.W. Raudenbush. 2009. "Exploiting Spatial Dependence to Improve Measurement of Neighborhood Social Processes." *Sociological Methodology* 39 (1): 151–83.

Sharkey, P. 2008. "The Intergenerational Transmission of Context." *American Journal of Sociology* 113 (4): 931–69.

Sharkey, P. 2012. "An Alternative Approach to Addressing Selection Into and Out Of Social Settings: Neighborhood Change and African American Children's Economic Outcomes." *Sociological Methods and Research* 41 (2): 251–93.

Sparks, A.L., N. Bania, and L. Leete. 2011. "Comparative Approaches to Measuring Food Access in Urban Areas." *Urban Studies* 48 (8): 1715–37.

Spielman, S.E., and E. Yoo. 2009. "The Spatial Dimensions of Neighborhood Effects." *Social Science and Medicine* 68 (6): 1098–1105.

Torrens, P.M. 2010. "Geography and Computational Social Science." *GeoJournal* 75: 133–48.

Weisburd, D. 2005. "Hot Spots Policing Experiments and Criminal Justice Research: Lessons from the Field." *Annals of the American Academy of Political and Social Science* 599 (1): 220–45.

Weiss, C.C., M. Purciel, M. Bader, J.W. Quinn, G. Lovasi, K.M. Neckerman, and A.G. Rundle. 2011. "Reconsidering Access: Park Facilities and Neighborhood Disamenities in New York City." *Journal of Urban Health* 88 (2): 297–310.

ESSAY

# The Dynamic Neighborhood Taxonomy

## *An Innovative Typology for Community and Economic Development*

Riccardo Bodini

The past decade has witnessed a renewed interest in urban communities: whether it was businesses seeking new investment opportunities, community-based organizations working to improve their neighborhoods, or government agencies reshaping their urban policy agenda, actors in the public, private, and nonprofit sectors have been seeking better ways to understand how to tap the assets of urban neighborhoods. In the wake of the economic crisis, this need has become more urgent: the problems have grown while the available resources have dwindled, placing a premium on targeted and strategic investment.

This type of strategic intervention primarily entails being able to target and tailor investments to the types of neighborhoods where they are most likely to be effective, which in turn requires accurate and timely information. In particular, it requires good and up-to-date data even for small levels of geography, sound knowledge on the dynamics of neighborhood change (including what factors can lead to neighborhood improvement or deterioration), and effective tools that can enable practitioners, investors, and service providers to make informed decisions on what actions to take in particular places.

Developing this type of information is no easy task, as neighborhoods are complex and dynamic entities that are constantly evolving: they are complex because they are composed of many different elements (people, businesses and institutions, infrastructure, and housing stock) interacting with each other across several dimensions. Neighborhoods are dynamic because they are in constant motion: even stable neighborhoods are constantly renewing their population, business base, and housing stock. Moreover, these dynamics are determined by the operation of social, political and economic systems (such as housing and labor markets, social networks, and local governance) that go well beyond the neighborhood itself. Neighborhoods are also diverse: there are different types of neighborhoods (ranging from bedroom communities to "bohemian" areas to "ports of entry") characterized by their mix of people, businesses, real estate, and physical environment. Different types of neighborhoods are likely to perform different functions for their residents and play different roles within the regional economy in which they are embedded.

The ability to measure and analyze complex neighborhood dynamics starts with good data, which are increasingly available thanks to initiatives like the National Neighborhood Indicators Partnership (NNIP). For many large cities at least, neighborhood- or even parcel-level data can now be found for various real estate, business, and demographic indicators.

These data now make it possible to engage in more detailed analyses and more in-depth investigations of the drivers of neighborhood change. Building on better data and knowledge, it is also possible to build more powerful tools to guide neighborhood investment. The Dynamic Neighborhood Taxonomy project, described in this essay, was designed as a baseline research and development effort with the goal of advancing community and economic development in this direction.

## The Dynamic Neighborhood Taxonomy Project

Living Cities, a collaborative of 22 of the world's largest foundations and financial institutions, launched the Dynamic Neighborhood Taxonomy project (DNT) in 2007 as an ambitious, large-scale effort to begin developing a new generation of community development tools. The project was designed to improve our understanding of how neighborhoods operate, including how they change over time, what factors determine their success, and how these basic dynamics vary across different types of neighborhoods. More important, DNT aimed to fundamentally enhance the field's ongoing capacity to routinely, accurately, and more easily analyze the challenges and opportunities for development in particular places. Ultimately, the goal was to help businesses, investors, funders, governments, and community development practitioners better tailor and target their investments and interventions in neighborhoods.

The project was conducted by RW Ventures, LLC,[1] in four cities (Chicago, Cleveland, Dallas, and Seattle) in partnership with numerous local and national organizations. The work was structured in four components: data collection on a wide variety of indicators, capturing the key dimensions of neighborhood change for every neighborhood in the four sample cities over at least 10 years; a descriptive analysis of how neighborhoods changed over this period; a series of regression models investigating the drivers of neighborhood change; and a typology of neighborhoods to enable investigating how patterns and drivers of change vary by neighborhood type.

In order to measure neighborhood change, the DNT project sought to develop a metric that would capture how current and potential residents value a community. In economic terms, this evaluation is reflected in the demand for housing in a neighborhood, and it can be measured in changes in housing values once we control for changes in the quality of the housing stock. Controlling for quality is important because the price of a house reflects the qualities of the structure itself (size, construction quality, number of bathrooms, and so forth) in addition to the desirability of its location. By holding quality constant, it is possible to estimate the change in price that can be attributed to a change in the desirability of the neighborhood and its amenities, rather than to changes in the characteristics of the house.

To measure quality-adjusted change in housing values, the DNT project developed a repeat sales index (which measures appreciation based on the repeated sales of the same house) at the census tract level. This index, which proved a powerful tool in and of itself, used a cutting-edge methodology to obtain reliable estimates of property appreciation at the census tract level while mirroring closely the reality of the housing market.[2]

In addition to changes in this repeat sales index, the project analyzed trends in median house prices as well as in the quantity of housing available in each census tract, in order to get a complete picture of the ways neighborhoods changed between 1990 and 2006.

These three metrics (the DNT Repeat Sales Index, change in median housing values, and change in housing quantity) were also used as dependent variables in a series of regression models designed to estimate the effect of various types of neighborhood amenities (including such things as retail and services, access to transit, and development interventions like Tax Increment Financing districts and Low Income Housing Tax Credit projects) on neighborhood change.

Building on this work and using a sophisticated cluster analysis technique, the project then constructed a typology of neighborhoods by grouping neighborhoods that were similar along key dimensions (such as demographics, business composition, and built environment), including in particular the factors that proved most important in the first phases of the analysis (that is, the strongest predictors of change in the dependent variables used by the project). The typology was designed to enable identifying comparable neighborhoods, facilitate peer analysis, and help anticipate and manage neighborhood change.

This body of work generated three kinds of outcomes for urban economic development: new knowledge about how neighborhoods evolve over time and what drives their evolution, new tools for neighborhood analysis and investment, and new capacity to take this kind of work further and apply it to other places and subject areas. All these outcomes, and their possible applications to community and economic development, are presented in the project's final report (Weissbourd, Bodini, and He 2009). This essay focuses on the DNT typology, which incorporates many lessons of the project and is a particularly innovative tool for neighborhood analysis. For present purposes, particular attention is paid to the methodology, structure, and applications of the typology, while the detailed outputs and findings it generated are described in depth in the DNT project's final report.

## The DNT Neighborhood Typology

Typologies help make complex issues manageable: by grouping communities that are similar along key dimensions, typologies help reduce complex and multidimensional entities such as neighborhoods to a few fundamental types, enabling meaningful comparisons, facilitating analysis, and revealing existing patterns. As it pertains to economic development practice, a neighborhood typology can be useful in at least three respects: it can help tailor strategies and interventions to the specific characteristics of particular neighborhoods, it can enable benchmarking the performance of each neighborhood and comparing it to its peers, and it can facilitate impact analysis by identifying comparable neighborhoods.

It should come as no surprise, then, that many different neighborhood typologies have been constructed over the years. A particularly good example is the one developed by The Reinvestment Fund (TRF), which constructed a typology of Philadelphia neighborhoods based on their real estate market characteristics, linking different types of neighborhoods to different types of housing interventions.[3] This typology reduced data on hundreds of thousands of properties to a manageable number of neighborhood types and helped Philadelphia's government prioritize interventions and better target its resources.

Several marketing and data companies have created neighborhood typologies of some sort by developing household segmentations based on consumer patterns, for the purposes of targeting product marketing and store locations. The PRIZM segmentation developed by Claritas, Inc., for example, defines the US market via 66 lifestyle groups characterized by different spending patterns. It then classifies neighborhoods based on their composition in terms of these segments. Academics and researchers also have developed numerous typologies of neighborhoods over the years, either as descriptive exercises or for analysis of particular phenomena.[4]

In fact, since the output of a typology depends entirely on what factors are used as inputs for the clustering algorithm, an infinite number of neighborhood typologies can be created, and none of them is necessarily more right or wrong than the others. Rather, typologies can only be evaluated in terms of how useful they are for the purposes for which they were developed. In this respect, existing neighborhood typologies have useful applications, but for various reasons none of them addressed the broader economic development issues tackled by the DNT project: they are often local (as in the case of the housing typology in Philadelphia) or based on a particular aspect of neighborhoods because they were designed to address a specific issue (consumer preferences in the case of PRIZM, housing investment in the case of TRF, and so forth). Many other existing typologies often end up being simpler descriptive exercises; they tend to present a static picture and are not grounded in an analysis of neighborhood dynamics and what drives neighborhood change.

The typology presented here was designed to help inform a broad range of community and economic development interventions, building upon the neighborhood analysis conducted by the DNT project. As such, it has some distinctive features that differentiate it from other neighborhood typologies. For example, it is multidimensional and grounded in the DNT project analysis of patterns and drivers of neighborhood change. By incorporating many of the factors that proved to make the most difference to the economic performance of neighborhoods, the neighborhood types can reveal something about the challenges and opportunities in each place.

Moreover, the DNT typology not only classifies neighborhoods based on their type, but also reveals all the connections between different types, in a hierarchical structure that goes from the broadest possible class to the narrowest grouping of neighborhoods that are most similar to each other. In practice, this means that the typology can be used to classify neighborhoods broadly and to identify, for any given neighborhood, its closest peers. Perhaps most important, this typology is dynamic: it incorporates the findings on patterns and drivers of change, and it shows how neighborhood types can transition to other types over time. Therefore, it can be used to help understand what should be expected in any given neighborhood and what interventions might be most appropriate.[5]

## Methodology

The starting point in building a typology is selecting the factors to use as determinants of neighborhood type. Drawing from the large number of variables at its disposal, the DNT project experimented with multiple combinations of variables and clustering methodologies in order to develop a typology that would help target and prioritize economic development interventions. As in the rest of the DNT project, the unit of analysis was the census tract.

The project started by building preliminary typologies using several different methodologies, including regression trees, expectation-maximization algorithms, and *k*-means and hierarchical clustering.[6] After evaluating each methodology, the project identified hybrid hierarchical clustering as the ideal method for the final version of the typology (Chipman and Tibshirani 2006). This methodology combines some of the advantages of *k*-means clustering with the hierarchical structure that can be obtained by hierarchical clustering. In other words, this method makes it possible to identify stable broad neighborhood types and at the same time drill down to very detailed subclusters within each type. This methodology also makes it possible to look at neighborhood types from the bottom up: for any given census tract, the clustering can identify which individual tracts are most similar along the selected dimensions, which can be especially useful for peer analysis for particular neighborhoods.

After settling on a methodology, the project experimented with different combinations of variables, testing the results based on statistical criteria (such as the importance in the models identifying the drivers of neighborhood change, the relevance to defining neighborhood types, and the reliability and availability of the data) as well as feedback from local partners in each of the four cities. This feedback helped evaluate typology outputs based on how well they seemed to group neighborhoods that were perceived as fundamentally similar by the people who knew them best.

The resulting neighborhood typology is based on variables that measure two key dimensions of neighborhoods: (1) the characteristics of the built environment (including, for example, characteristics of the housing stock, land use patterns, and business presence); and (2) the characteristics of the people who live there (such as income, age, household structure, and mobility). Each dimension was given equal weight in the typology to ensure that both categories are equally important in determining the final neighborhood groupings regardless of how many variables were chosen under either category.

Before selecting actual variables, the project identified within each dimension a set of subcategories that appeared to define neighborhood type in the various typology iterations. For example, within the "people" dimension, "age structure" is a subcategory that contains all the different age group variables. Again, each subcategory was given equal weight within its dimension.[7]

This layered structure—composed of dimensions, subcategories, and variables—helps ensure that each factor is given appropriate weight in determining neighborhood type regardless of how many variables are used to measure it. For example, if the age structure subcategory included four age brackets (and, thus, four different variables) and income diversity included one variable, a non-weighted version of the typology would give age structure four times more importance than income diversity. By weighting the variables within each subcategory and dimension, all subcategories become equally important. A table with the final hierarchy of categories, subcategories and variables, along with their final weights, can be found in the appendices to the DNT final report.[8]

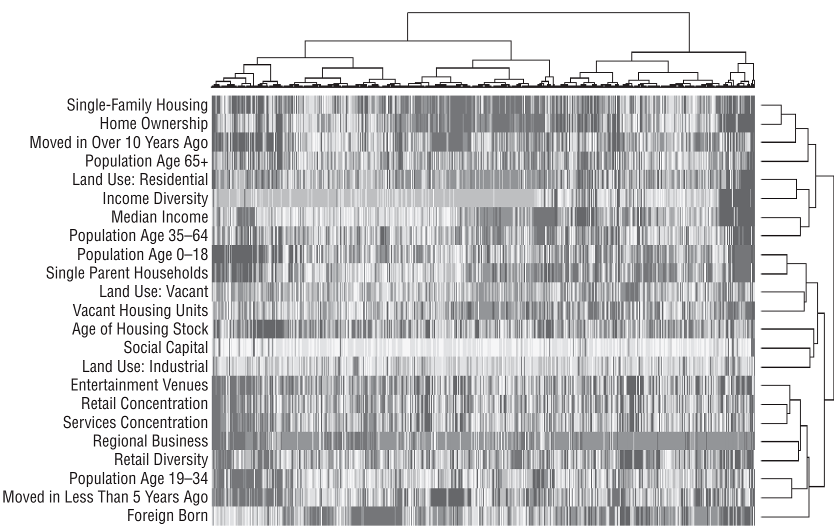
As mentioned above, this typology sought to identify not only distinct neighborhood types, but also how each type changes over time. To this end, each neighborhood (i.e., census tract) was included in the typology twice: once based on its characteristics in 1990 and once based on its characteristics in 2000.[9] This is a new, very important feature of this typology; it can reveal a great deal about neighborhoods' dynamics of change, and it helps identify which types are more likely to change over time and transition to other types.

## Structure

Figure 7.1.1 summarizes the structure of the typology. The figure is a heat map, a technique first developed as a DNA mapping application in the genome project, where it was used to examine how different groups of genes correlate with various physical traits. The same technique is applied here to display how neighborhoods are grouped together according to their score on the 23 different variables selected as key determinants of neighborhood type. What makes this application particularly useful here is that it works well when multiple factors interact in complex ways, as in the case of the various dimensions that determine neighborhood types.

The map can be interpreted as a grid in which each column is a census tract and each row is a variable. Although shown in black and white here, the heat map is presented in shades of blue and red in the DNT final report. The score of each tract on the variables listed to the left of

**Figure 7.1.1.** Typology Heat Map



the chart is represented in our source analysis by degrees of color, from dark red (very low) to dark blue (very high). The neighborhood types are created by grouping neighborhoods that tend to have similar scores on the same variables, which can be viewed in our report by the concentrations of red and blue cells on the map. For instance, a blue area on the bottom left of the map identifies a group of neighborhoods with a high percentage of young adults, high income levels, and a high concentration of retail, services, and entertainment venues. Similarly, a blue area at the top of the central section of the map identifies a group of mostly residential neighborhoods characterized by older residents, high homeownership rates, and prevalence of single-family homes.

The figure contains two additional important pieces of information. The first is that neighborhoods that are closer together in the chart are more similar than neighborhoods that are further apart. These relationships are summarized in the tree structure on the top of the figure. In this sense, the typology generated by the DNT project is actually a taxonomy of neighborhoods: indeed, it works just like a taxonomy of living organisms in biology, which organizes all forms of life in a hierarchical structure that goes from the broadest grouping of kingdom to phylum, then class, and so forth all the way down to species.

This means that the typology can be used from the top down as well as from the bottom up. In other words, we can start with the broadest possible grouping of neighborhoods and further refine our types as we move down the tree. Or, we can start with a particular neighborhood and identify which other neighborhoods are most similar to it. The top-down approach is useful to surface general findings regarding a particular neighborhood type, such as its likelihood of undergoing particular changes or the interventions most likely to make a difference. The bottom-up approach, on the other hand, can be used to see how a particular neighborhood is doing relative to its peers, or to evaluate the impact of a specific intervention.

The second piece of information is that the same hierarchical structure is applied to the variables used to build the typology. Therefore, variables that are closer together in the figure tend to be correlated to each other and have similarly high or low values in the same neighborhoods, revealing how different factors combine to determine neighborhood types. In particular, the tree to the right of the map shows there are three main groups of variables. The first group has to do with the stability of the neighborhood and its housing stock, and it includes the percentage of single-family homes, homeownership rates, residential land use, and median income. The second group includes a number of indicators typically associated with neighborhood distress: vacancy rates (both vacant land and vacant housing units), percentage of single-parent households, social capital, and industrial land use. The third group has to do with the concentration of retail and services in the neighborhood, indicators that also tend to be associated with the presence of a younger and more mobile population.

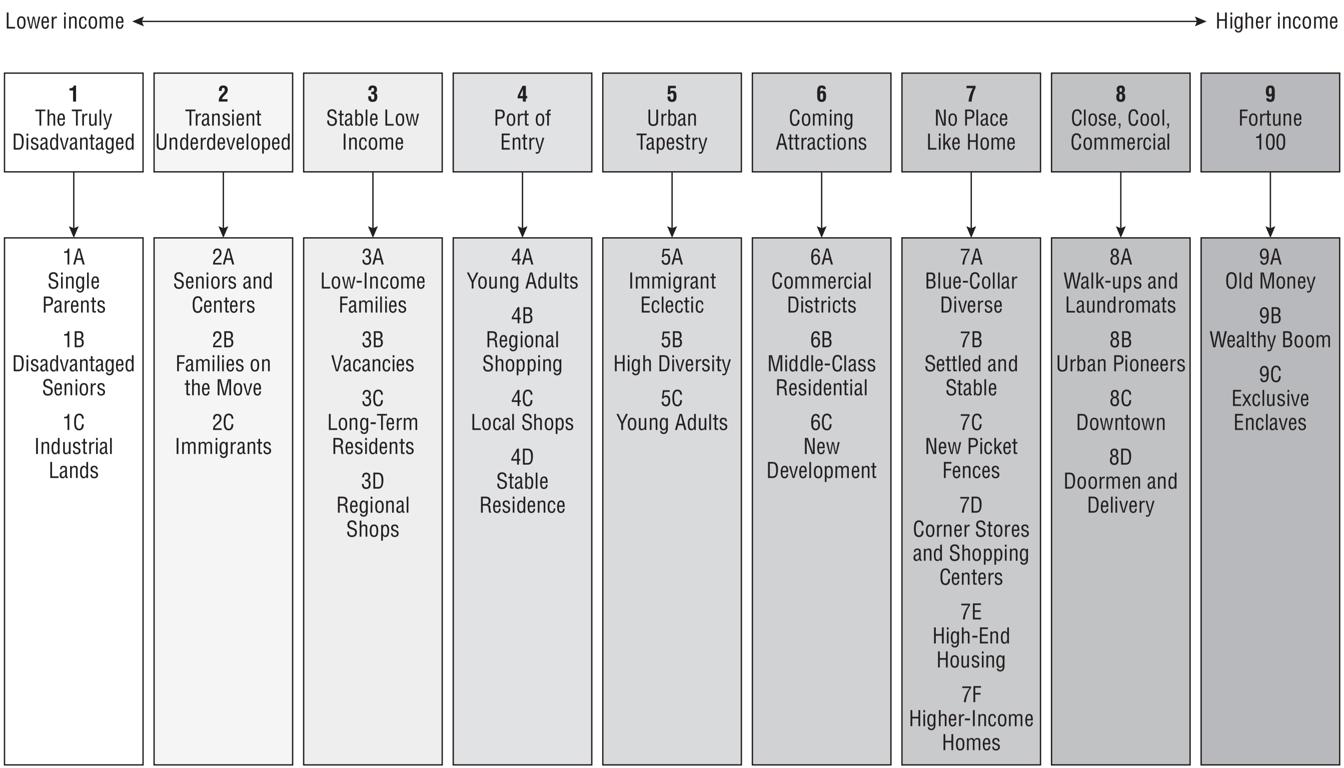
The hierarchical structure of the typology yields a potentially very large number of neighborhood types, as we can keep refining each grouping until we reach the individual neighborhoods at the bottom of the tree. To make this information useful and accessible, however, a manageable number of distinct neighborhood types should be identified, while at the same time preserving enough differentiation between types to see real differences in characteristics and drivers of change. To achieve this balance, the project focused on two layers of the typology, deriving nine broad neighborhood types that were then further divided into several distinct subtypes.[10]

Figure 7.1.2 summarizes these two key layers of the typology, displaying the nine types and 33 subtypes of neighborhoods identified with this system. The broad neighborhood types are ordered based on their median income and numbered from 1 to 9. Within each type, the subtypes are ordered based on their median income and assigned a letter. Therefore, type 1-A (single parents) is the lowest income segment, while type 9-C (exclusive enclaves) is the wealthiest.[11]

While only 23 variables were used to construct the typology, many more variables can be used to profile each neighborhood type. For instance, it is possible to describe each type by its location, racial composition, residents' occupations, or even foreclosure or crime rates, even though none of these factors were used to define the type in the first place. By and large, the types are well differentiated based on these other descriptive features as well, lending validity to the final classification.[12] For instance, while race was not included as a defining variable, neighborhoods with a distinct racial makeup tended to fall into several distinct neighborhood types.

Similarly, while the typology was constructed by pooling all neighborhoods in the four cities, not all types are found everywhere. For example, none of the three poorest types are found in Seattle, where incomes are generally higher than in the other three cities. Conversely, Coming Attractions neighborhoods are found primarily in Dallas and Seattle; very few neighborhoods in Chicago, and none in Cleveland, match this profile.

**Figure 7.1.2.** Neighborhood Types and Subtypes



This typology also enables us to see whether different types tend to change over time, since it is constructed using data from 1990 and 2000. This information is summarized in the transition matrix reported in figure 7.1.3, which shows what percentage of neighborhoods in each type remained the same type 10 years later (the darker the color, the higher the percentage), and, if a neighborhood changed type, what other type of neighborhood it usually became. The matrix shows, for instance, how most neighborhoods tend to not change their type within these 10 years, though there are significant differences across neighborhood types. In particular, the low-income segments tend to change type more often than the higher-income segments: approximately 65 percent of type-1 neighborhoods were a different type 10 years later, while less than 20 percent of the neighborhoods in the three wealthiest segments changed type between 1990 and 2000. The information yielded by this transition matrix was included in the profiles for each neighborhood type.

**Figure 7.1.3.** Transition Matrix

| 1990 \ 2000 | Type 1 | Type 2 | Type 3 | Type 4 | Type 5 | Type 6 | Type 7 | Type 8 | Type 9 | Count in 1990 |
|---|---|---|---|---|---|---|---|---|---|---|
| Count in 2000 | 47 | 153 | 301 | 169 | 193 | 69 | 344 | 128 | 55 | |
| Type 1 | 35.8 | 40.8 | 19.2 | 0.8 | | 0.8 | | 1.7 | 0.8 | 120 |
| Type 2 | | 45.5 | 16.7 | 16.7 | 11.1 | 4 | 0.5 | 5.1 | 0.5 | 198 |
| Type 3 | 1.2 | 2 | 89.8 | 0.4 | 2.5 | | 4.1 | | | 244 |
| Type 4 | 0.9 | 3.7 | | 74.1 | 17.6 | | | 3.7 | | 108 |
| Type 5 | | 1.7 | 7.2 | 21.1 | 58.9 | | 2.8 | 7.8 | 0.6 | 180 |
| Type 6 | | 1.2 | | 9.5 | 13.1 | 61.9 | 3.6 | 9.5 | 1.2 | 84 |
| Type 7 | | | 3.6 | 2.2 | 6.6 | 0.3 | 86.3 | 0.5 | 0.5 | 366 |
| Type 8 | | | | | 4.9 | 3.9 | 1 | 85.3 | 4.9 | 102 |
| Type 9 | | 1.8 | | | | 5.3 | 14 | 1.8 | 77.2 | 57 |

## Sample Output

While in many ways this typology is still just a prototype, it works fairly well for the purposes for which it was created: it identifies distinct neighborhood types that present specific challenges and opportunities, it reveals important facts about the patterns and drivers of change of each type, and it enables the user to identify comparable neighborhoods along the dimensions that matter most for community and economic development. A detailed analysis of what can be learned about neighborhoods by using this typology can be found in the DNT final report and would be beyond the scope of this essay. Nevertheless, it is useful here to include at least one example of the information it contains.

In general, for each neighborhood type identified by the typology, the project developed a detailed profile, designed to summarize the information most relevant for community and economic development. These profiles include a high-level description of the neighborhood type, followed by a more technical profile with additional details on the factors that characterize that type of neighborhood. The description also includes observations on the neighborhood's dynamics of change (including key implications for economic development) and a summary of the growth trends for the two key metrics used as indicators of neighborhood improvement (Repeat Sales Index, or RSI, and change in quantity of housing units) for the neighborhoods that did not change type between 1990 and 2000. Finally, the profile includes a chart showing the distribution of the cluster across cities and summaries of each subtype into which the broad type can be divided.

The profile of type 4 (Port of Entry) neighborhoods is shown in table 7.1.1 as an example of the typology output. The profiles for all neighborhood types can be found in the DNT final report.

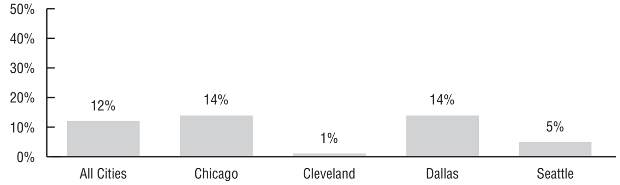
## Applying the Typology

The first and most immediate use for the typology is to benchmark neighborhood performance. Identifying neighborhoods that have the same overall characteristics ensures that comparisons actually make sense. For instance, a neighborhood can use the typology to compare its growth rates to those of the other neighborhoods in the same subtype, and see whether it is doing as well as could be expected. As an example, let's take a census tract in the North Lawndale neighborhood on the West Side of Chicago. This tract belongs to type 2-A (Seniors and

**Table 7.1.1.** Sample Type Profile

| | **Type 4: Port of Entry** |
|---|---|
| **Highlights** | Blocks animated by a variety of businesses and residents' native languages make up the neighborhoods of Cluster 4. Most type 4 neighborhoods have a Hispanic majority, though these communities can also be enclaves of Asian and European immigrants. Almost half of neighborhood residents were born outside the United States, and many are raising families in these parts of the city with little crime and well-used space. Many residents move from their homes—few of which are single-family dwellings—after a few years but may stay in the neighborhood. Residents have slightly lower-than-average incomes, but unemployment is less than 10 percent, and two parents are present in most households with children. |
| **Detailed profile** | Neighborhoods in this cluster represent the bulk of the "immigrant communities" in the typology, with 45 percent of their population foreign-born. Although most of these neighborhoods are primarily Hispanic, a few are majority non-Hispanic white or Asian (particularly in Seattle).<br><br>Cluster 4 lies on the line between low- and moderate-income clusters ($32,000 household income on average), but its socioeconomic indicators are more similar to the mid- to higher-income clusters than to the lower-income groups. In particular, these neighborhoods are characterized by lower unemployment rates, lower percentages of single-parent households, and greater income diversity. Resident mobility is relatively high, consistent with the "port of entry" character of these communities.<br><br>Employment in these neighborhoods tends to be concentrated in a few specific occupations, more so than in other neighborhood types. About 24 percent of adults in the labor force are employed in production and transport occupations, and 12 percent of residents are employed in the construction sector, both of which are the highest rates among all clusters. Conversely, the proportion of residents in professional, sales and office occupations are considerably lower than in the other mid- to high-income clusters. This is consistent with the fact that, on average, 47 percent of adults do not have a high school diploma.<br><br>Business presence is among the highest of all clusters. However, the types of businesses that characterize these communities vary greatly within this cluster; some neighborhoods have a greater presence of local shops, while others have a greater concentration of large business establishments. |

**Table 7.1.1.** *(Continued)*

| Type 4: Port of Entry | |
|---|---|
| **Dynamics of change** | Type 4 is a moderately stable type: 74 percent of these neighborhoods in 1990 remained in the same cluster by 2000. At the same time, however, about 18 percent transitioned into type 5 (Urban Tapestry), 4 percent transitioned into type 2 (Transient Underdeveloped), and 4 percent transitioned into type 8 (Close, Cool, and Commercial). Different subtypes are more likely to transition to different clusters, though, as discussed in more detail below.<br><br>Gentrification appears to be a driving force behind the transition of some of these neighborhoods to higher-income clusters. Features such as the cluster's overall proximity to downtown, lower housing values, and retail amenities may be contributing factors. In some instances this change is more gradual, and these neighborhoods transition to Urban Tapestry communities; in other cases the process is more rapid, and these communities transition to Close, Cool and Commercial neighborhoods.<br><br>In neighborhoods that remain largely immigrant communities, an important driver of improvement is the presence of employment opportunities nearby, as proximity to jobs is particularly important for this population. |
| **Growth trends** | On average, the RSI had the third-fastest growth rate between 1990 and 2000, increasing 17 percentage points faster than the city as a whole, and the fourth-fastest growth rate between 2000 and 2006, increasing 5 percentage points over the citywide rate.<br><br>Growth rates in the quantity of housing units were slightly below average between 1990 and 2000 (–3 percent), and above average between 2000 and 2005 (3 percent). |
| **Presence by city (2000)** |  |



*(continued)*

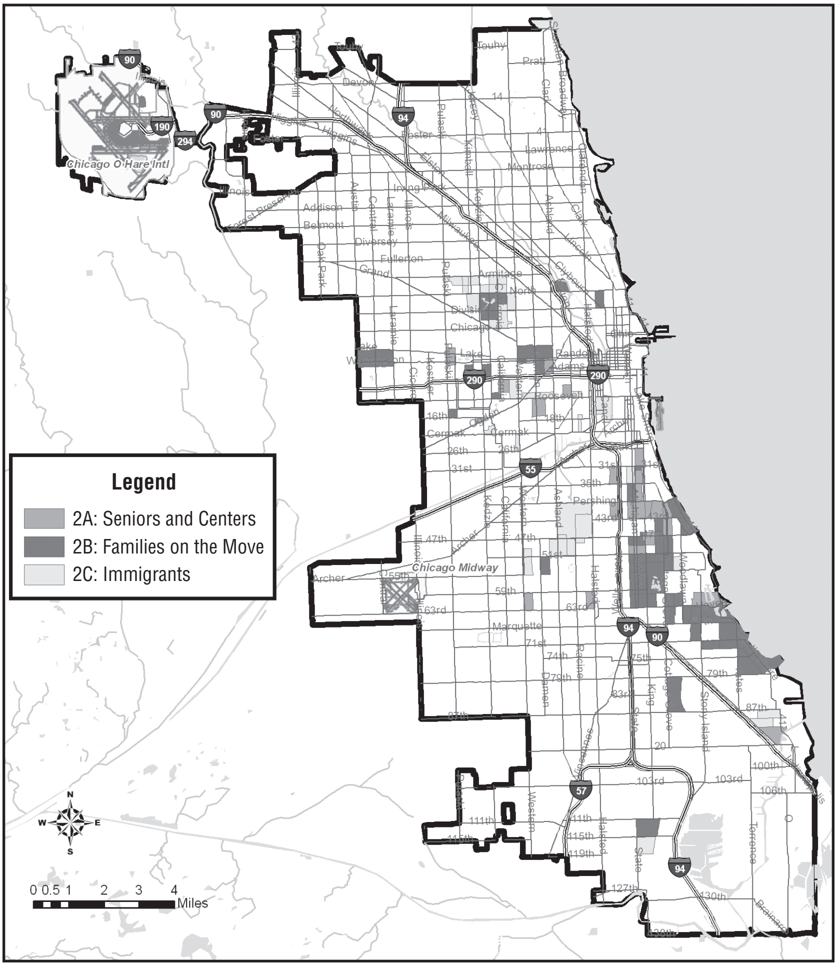
**Table 7.1.1.** *(Continued)*

| | Port of Entry: Subtypes |
|---|---|
| **Type 4-A: Young Adults** | Neighborhoods in type 4-A are inhabited by a younger, more mobile population. About 78 percent of the population has changed residences over the past five years, compared with 60 percent for the overall cluster, while only 12 percent have lived in the same home for over 10 years. Residents are also less likely to have a high school diploma, and about 39 percent have jobs in construction, production, or transportation occupations. An above-average percentage of the land is vacant, and business presence is lower than the average for this cluster, suggesting that as a whole, this subtype is less developed than the other Port of Entry neighborhoods. Despite the high mobility of its residents, this group does not often change type. |
| **Type 4-B: Regional Shopping** | Neighborhoods in type 4-B are characterized by the presence of larger businesses and have the highest concentration of retail within the Port of Entry type. These neighborhoods are more likely to transition to higher income clusters, perhaps due to lower crime rates, a high concentration of retail amenities and entertainment venues, and proximity to downtown. |
| **Type 4-C: Local Shops** | These communities are characterized by the presence of smaller, local businesses. As in the case of type 4-B, these neighborhoods are more likely to transition to higher income clusters. |
| **Type 4-D: Stable Residents** | Type 4-D typically includes older, more established communities. About 30 percent of the housing stock is composed of single-family homes, and 39 percent of households own their unit. Median incomes are higher, and residents enjoy the greatest diversity of business types within this cluster. Consistent with the more stable character of these communities, this subtype is much less likely to transition to other neighborhood types. |

Centers). Using the typology, peer neighborhoods can be identified across the entire city (figure 7.1.4). Based on the RSI,[13] the growth rate in this tract between 1990 and 2006 was 33 percent, which is much lower than the average for its type (approximately 100 percent over the same period), revealing that this neighborhood is not doing as well as its peers.

The typology can also be used to go much further than this initial diagnostic. In fact, it points at some factors that might be particularly

**Figure 7.1.4.** Location of Neighborhoods in Transient Underdeveloped Category (Type 2)




important for each type, and it enables us to compare a neighborhood to its peers with on those factors, in order to identify priority areas for development interventions.

For instance, having identified a group of peer neighborhoods for the census tract in North Lawndale, we can use the typology to see how this tract is doing on some key drivers of neighborhood change for Stable

**Table 7.1.2.** North Lawndale, In-Type Comparison

| | Tract 291800 | Type 2-A, Seniors and Centers |
|---|---|---|
| Change in value (RSI, 1990–2006) | 33% | 100% |
| Median income | $18,560 | $17,000 |
| Vacant units | 29.2% | 19.0% |
| Social capital | 3.72 | 3.95 |
| Unemployment rate | 37.0% | 20.9% |
| Turnover (% moved in past five years) | 55.2% | 30.8% |
| Educational attainment: no high school diploma | 69.1% | 42.6% |

Low-Income neighborhoods (as identified by the models developed in the other phases of the DNT project). An initial comparison (summarized in table 7.1.2) reveals that this tract is actually trailing its peers on a few important dimensions, including employment, educational attainment, and, to a lesser extent, social capital[14]—all of which are drivers of improvement in place. This information can then be used to prioritize interventions and address these particular areas.

The typology can also be used to go one step further and find out how similar neighborhoods have been dealing with these issues. In particular, the typology can be used from the bottom up to find the neighborhoods that are most similar to a particular place (in the case of the North Lawndale neighborhood discussed above, several neighborhoods on the West and South Side of Chicago, as well as some communities in South Dallas), see which ones have successfully dealt with the same development issues (in this case, unemployment and educational attainment), and identify the interventions that can best be adapted and applied.[15]

An additional important use for the typology is that it can help analysts think strategically about the trajectory in which a neighborhood is headed and what kind of place its residents would like it to be. Consider, for instance, neighborhoods in type 6 (Coming Attractions). Based on the dynamics of change identified using the transition matrix, we know that these areas can evolve in several directions and become very different types of places. By comparing the characteristics of these different types, we can identify areas of intervention that would help push the neighborhood in the desired direction. For instance, if the residents of a type-6 neighborhood determined that they would like their community to evolve in the direction of the Close, Cool, and Commercial neighborhood type, development interventions should focus more on retail development and other amenities that can attract the key demographics of that neighborhood type. Conversely, if the desired trajectory is more in the direction of type 7 (No Place Like Home), interventions should focus on housing, homeownership, and school quality.

Moreover, the typology can be used to take a more granular and nuanced look at a particular community, targeting interventions to different parts of the area and anticipating the changes that might lie ahead. Take, for instance, the community area of Chicago Lawn on the southwest side of Chicago, highlighted in figure 7.1.5. While people think of Chicago Lawn as one neighborhood, as many as five distinct neighborhood types and seven distinct subtypes are within its boundaries, including No Place Like Home to the south, Stable Low-Income to the east, Urban Tapestry in the middle, and Port of Entry to the northwest.

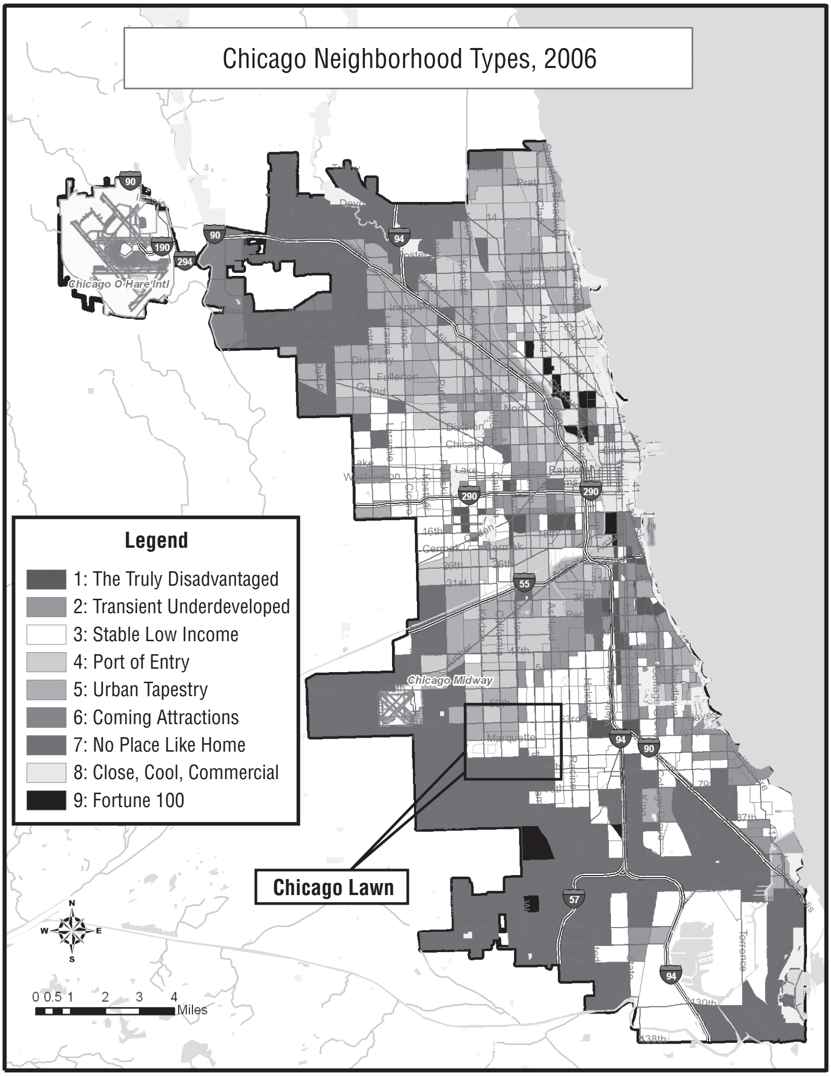
Additional information can be gleaned by applying the transition matrix that shows how each tract has evolved over time. The pattern of transitions between 1990 and 2000 points to significant changes taking place in this neighborhood, particularly due to an influx of lower-income and foreign-born households (in the eastern and northwestern portions of the community, respectively). These trends are likely to continue (partly owing to gentrification and displacement in communities closer to downtown) and to cause the Urban Tapestry portion of this community to take on more characteristics of Port of Entry and Stable Low-Income neighborhood types. The information contained in the typology on these different types can also help prioritize interventions and would suggest focusing in particular on issues related to safety, foreclosures, and homeownership in the Stable Low-Income portions of the neighborhood and access to jobs in the Port of Entry sections.

## Limitations and Opportunities for Improvement

This typology should be thought of as a prototype that needs to be tested and refined. Like much of the work done by the Dynamic Neighborhood Taxonomy project, it is intended as a foundation that can be expanded and improved over time, in several ways.

First, every time the typology is used and applied to a particular place within the four cities for which it was developed, it produces feedback that can be used to increase the typology's accuracy. Indeed, when the

**Figure 7.1.5.** Chicago Neighborhood Types and the Chicago Lawn Neighborhood



information contained in the typology is compared to the reality on the ground, inaccuracies and imperfections will undoubtedly surface. This information can then be used to understand the instances in which the data are misleading, fine-tune the profiles, and make them more accurate and useful.

The typology could also be expanded in several ways. It would be relatively easy, for instance, to develop a next layer of neighborhood subtypes. This information is already embedded in the hierarchical structure of the typology, and it needs to be extracted and made accessible. Doing so would yield a more detailed picture of each type and enable a more granular analysis of particular places.

An additional next step would be to update the data and track changes in neighborhood type over a longer period. This would serve two important purposes: it would ensure that the typology is always relevant and up to date, and it would reveal additional information about patterns of change and transitions between neighborhood types that could then be incorporated in the type profiles and enhance their value.

Finally, and perhaps most important, the typology could be expanded to include neighborhoods in other cities. This step would probably be the most valuable in order to increase the applicability and usefulness of the typology. Every time a new neighborhood is added, it enriches the base of information used for the typology and increases its accuracy, both in profiling each type and in revealing useful information regarding their patterns and drivers of change. Moreover, while the four cities selected for the analysis encompass a wide variety of neighborhood types, many neighborhoods in other cities certainly would not fit neatly in any one of them, and might in fact represent entirely new types that are not included in this typology. By adding them to the sample, we would expand the scope of the typology and ensure that it is broadly applicable beyond the four cities for which it was originally constructed.

These improvements would successfully build on the innovative approach of this typology while addressing its current structural limitations (related primarily to data availability and timeliness), but the main challenge remains making tools like this routinely available to those who could use them—that is, community development practitioners and policymakers. This issue, of course, transcends this specific tool and its applications, but it should be a chief concern for all of us who seek to develop better data and methodologies for neighborhood analysis and development.

The richness of available data and analytics has enabled us to reach an unprecedented level of sophistication and accuracy in measuring and understanding neighborhood dynamics. Still, there is much work to do in order to bridge the worlds of data and analysis and the world of everyday community and economic development practice. The progress being made in developing interactive, web-based platforms that integrate rich datasets and GIS software certainly goes in this direction. However, up to this point these platforms are mostly used to display data for descriptive purposes rather than to deploy new analytic tools like the DNT neighborhood typology. In this sense, the new frontier in this line of work remains the development of a user interface that can enable people who do not have a background in data analysis and statistics to access, understand, and deploy in their own work the information most relevant to them.

## NOTES

1. RW Ventures, LLC, is an economic development firm focusing on market-based strategies for regional and community economic development. More information on RW Ventures and the Dynamic Neighborhood Taxonomy Project can be found on the firm's website, www.rw-ventures.com.

2. The index was developed based on the work of McMillen and Dombrow (2001) and reflects significant improvements. The procedures and methodology developed to estimate this index are explained in more detail in appendix D of the DNT final report (Weissbourd, Bodini, and He 2009). The report also contains full color versions of the figures and tables presented in this essay.

3. See http://www.trfund.com/market-value-analysis-philadelphia/.

4. See, for example, McWayne et al. (2007).

5. Ultimately, a detailed analysis of the drivers of change for each neighborhood type would be needed in order to complete this picture. The typology presented here is an important step in this direction.

6. Among other things, the project experimented with a methodology that would classify neighborhoods directly based on their drivers of change—that is, using the coefficients in the regression models to isolate groups of neighborhoods that would respond to the same type of interventions. While this approach would ensure that the typology could help target interventions to the types of neighborhoods where they are most likely to be effective, it could not be fully implemented owing to an insufficient sample size. If other cities were to be included in the analysis, expanding the number of neighborhoods in the sample, this approach could be resumed and would likely produce powerful results.

7. The only exception is income, which was weighed more heavily since it affects so many other neighborhood characteristics.

8. The appendix also includes a detailed description of all the variables, including their summary statistics.

9. The years 1990 and 2000 were selected for this exercise because the most data were available for these years. However, to make the results more applicable, when the typology is applied to a particular place, each neighborhood is assigned to a type based on the most current data available.

10. Different factors tend to define different layers of the taxonomy. At the highest level, a neighborhood's type appears to be defined primarily by its housing stock, the income of its residents, and the share of the population that is foreign-born. The next differentiation happens based on the age of the population (which is likely related to the preferences for different types of neighborhood amenities), land use patterns, and business presence.

11. Income was chosen as a key dimension for two reasons: it is a very important outcome from an economic development standpoint, and it plays a very important role in determining neighborhood type.

12. The validity of a typology cannot be tested based on differences in the variables used for the clustering, as those will, by definition, differ more across types than within types. However, it is possible to test the validity of a typology based on differences in variables not used for the clustering. If those variables are well differentiated across types (as was the case for this typology), this indicates that the clustering surfaced truly distinct neighborhood types.

13. As described in the previous section, the RSI was not a defining variable in the typology. However, the RSI was included in the neighborhood profiles, and it was used to validate the typology results by verifying that neighborhoods grouped together by the typology also tended to have similar values in the RSI.

14. For the purposes of the DNT project, social capital was measured based on the presence in the tract of selected types of organizations, including civic and social associations, churches, and membership organizations. See the DNT final report for details.

15. These comparisons can also be drawn across time: since the typology includes 1990 observations, a neighborhood today could find other neighborhoods that were in the same situation 20 years ago, and see what those neighborhoods did and how they evolved.

## REFERENCES

Chipman, Hugh, and Robert Tibshirani. 2006. "Hybrid Hierarchical Clustering with Applications to Microarray Data." *Biostatistics* 7: 302–17.

McMillen, Daniel P., and Jonathan Dombrow. 2001. "A Flexible Fourier Approach to Repeat Sales Price Indexes." *Real Estate Economics* 29 (2): 207–25.

McWayne, Christine M., Paul A. McDermott, John W. Fantuzzo, and Dennis P. Culhane. 2007. "Employing Community Data to Investigate Social and Structural Dimensions of Urban Neighborhoods: An Early Childhood Education Example." *American Journal of Community Psychology* 39: 47–60.

Weissbourd, Robert, Riccardo Bodini, and Michael He. 2009. *Dynamic Neighborhoods: New Tools for Community and Economic Development.* Chicago: RW Ventures, LLC. Available at www.rw-ventures.com/publications/n_analysis.php.

ESSAY

# Neighborhoods, Neighborhood Effects, and Residential Mobility

## *A Holistic View and Future Directions*

George Galster and Lina Hedman

Though urban neighborhoods have been the object of long-standing scholarly attention, interest has increased dramatically over the last quarter century. Following the publication of Wilson's (1987) now classic analysis of the development of poverty areas and the new urban underclass, a growing number of scholars, scientific programs, foundations, planners, and community development practitioners have directed their attention to answering the two preeminent questions about urban neighborhoods: What makes them change physically, demographically, and socioeconomically? To what extent do they affect their residents, independently of individual characteristics and other forces?

In an effort to answer these questions, two substantial interdisciplinary literatures have developed. One has focused on neighborhood selection and sorting by income, ethnicity, and other household characteristics, building our understanding about why and where people move. The other has focused on neighborhood effects, building our understanding about the degree to which residential context exerts independent effects on a variety of outcomes for residents. Unfortunately, these two literatures have long remained artificially segregated from each other in conceptual and empirical terms.

Galster (2003), Doff (2010), and Hedman (2011) have provided distinctive but complementary treatises attempting to link these two literatures conceptually.[1] In summary, they argue that (1) individual characteristics affect what neighborhood is selected when an individual moves; (2) the neighborhood selected can subsequently affect some individual behaviors; (3) those behaviors, in turn, affect whether that individual remains in the current neighborhood and, if not, what different neighborhood will be selected; and (4) failure to recognize these interrelationships leads to biased statistical estimates both of neighborhood effects and determinants of mobility. The upshot of their arguments is that both residential mobility and neighborhood effects literatures would be enriched by a more holistic, unifying approach.

In this essay we build on this foregoing work to extend and illustrate such a holistic view in conceptual and empirical terms. We first advance a synthetic model that delineates the numerous interconnections between neighborhood conditions, neighborhood effects, and residential mobility. We then provide overviews of the salient elements of the neighborhood effects and the residential mobility literatures from the United States and Western Europe, focusing on evidence related to the mutually causal interconnections between them specified in our model. We proceed to explain why a more holistic view is crucially needed to advance the statistical modeling efforts within both literatures, focusing on avoidance of geographic selection and endogeneity biases. Next we review the scant literature that has taken important preliminary steps in this holistic direction. Finally, we conclude and suggest challenges for cutting-edge research related to neighborhood change, effects, and mobility.

## A Holistic View of Neighborhood, Neighborhood Effects, and Residential Mobility

Our thesis in this essay can be introduced succinctly through seven generic equations (or, perhaps more accurately, shorthand sentences) that delineate in simplified fashion the complex causal interrelationships among what the neighborhood is, what effects it may have on its residents in aggregate, and how its residential composition may change in aggregate through in- and out-mobility processes. For brevity of notation, let *j* represent an indicator designating a particular individual and *n* represent the neighborhood in which the individual resides during time period *t*. Endogenous elements to the system (i.e., aspects that are mutually causal during period *t*) are italicized for emphasis. The notation $f(\ldots)$ denotes "is a function of ...." For the moment we ignore the thorny question about what geographic entity delineates *n.* We posit that

1. *Neighborhood n Context during t = f(Socioeconomic and Demographic Characteristics and Behaviors of all j in n during t; Social Relationships among j in n during t;* Other Institutional, Political, Environmental, and Topographical Characteristics of n during t)
2. *Socioeconomic and Demographic Characteristics and Behaviors of all j in n during t = f(*Socioeconomic and Demographic Characteristics and Behaviors of all j in n at start of t; *Who Moves Into n during t; Who Moves Out of n during t; Neighborhood n Context during t; Duration of Exposure of j to n; Stability* of *Neighborhood n Context from prior periods through t*)
3. *Social Relationships among j in n during t = f(*Socioeconomic and Demographic Characteristics and Behaviors of all j in n at start of t; *Who Moves Into n during t; Who Moves Out of n during t; Neighborhood n Context during t; Stability of Neighborhood n Context from prior periods through t;* Social Institutional Infrastructure)
4. *Who Moves Into n during t = f(Socioeconomic and Demographic Characteristics and Behaviors of all j in n during t;* Other Institutional, Political, Environmental, and Topographical Characteristics of n during t; Socioeconomic and Demographic Characteristics and Behaviors of Other Neighborhoods besides n during t; Socioeconomic and Demographic Characteristics and Behaviors of all j in Metro Housing Market at Start of t; Other Structural Characteristics of Metro Area Transport Systems, Energy Costs, and so forth)
5. *Who Moves Out of n during t = f(Socioeconomic and Demographic Characteristics and Behaviors of all j in n during t;* Other Institutional, Political, Environmental, and Topographical Characteristics of n during t; Socioeconomic and Demographic Characteristics and Behaviors of Other Neighborhoods besides n during t; Other Structural Characteristics of Metro Area, Transport Systems, Energy Costs, and so forth)
6. *Duration of Exposure of j to n = f(Who Moved Into n during t; Who Moved Out of n during t*)
7. *Stability of Neighborhood n Context from prior periods through t = f(Who Moved Into and Out of n during t; Who Moved Into and Out of n prior to t;* Changes in Other Institutional, Political, Environmental, and Topographical Characteristics of n during t and prior to t)

These seven equations attempt to summarize the currently accepted scholarly wisdom about the factors affecting indicators and behaviors of interest. That is, equation 1 represents a definitional statement that the neighborhood is a multidimensional package of attributes, some created endogenously by residents, other created by exogenous institutional or political forces, others bequeathed by nature (Galster 2001). Equation 2 is partly a tautology that the aggregate profile of a neighborhood during a period will be determined by the starting population and mobility-produced changes in that population (including births and deaths). But it also includes the possibility that the character of the residents may be shaped endogenously by the neighborhood itself through one or more neighborhood effect mechanisms (Galster 2012). Equation 2 suggests that a given set of neighborhood characteristics will have differential power depending on the characteristics of the residents (such as age, gender, and ethnicity), but also by how long a person is exposed to the environment and how stable that context is over time (Galster 2012). Equation 3 posits that the nature and intensity of social relations in a neighborhood will be determined by the characteristics of the residents and the stability of those residents, as well as by the context established by aggregate neighborhood characteristics and behaviors and by social institutional infrastructure. Equations 4 and 5 constitute the residential mobility equations and are founded on the notion that neighborhoods compete for residents with other neighborhoods that offer closely substitutable residential packages (Galster 1987; Rothenberg et al. 1991). Current residents of a particular neighborhood evaluate its context in comparison to the contexts of what the household perceives as feasible alternative neighborhoods, with relative evaluations being shaped by individual characteristics as well as metrowide forces that can exogenously alter the hierarchy of neighborhoods according to expense and/or attractiveness. The current neighborhood social context may be influential in shaping the out-migration decisions of current residents, but it will typically be less important to potential in-movers because they cannot assess prospectively the social climate and/or accurately predict how they will respond to it once in residence. Existing social relations (i.e., having friends and family in a neighborhood) may, however, be important to in-movers (Hedman 2013).[2] Equations 6 and 7 complete the framework by noting how both the consistency of the residential context over time and the duration of the residents' exposure to it will be shaped by their mobility behavior, both individually and in aggregate.

The upshot of this formulation should be obvious. What neighborhood context is, how temporally stable it is, and what impact it has on its residents are influenced by the mobility behavior of its residents and others in the same metropolitan housing market. But the mobility behavior of its residents is influenced by what neighborhood context is, how temporally stable it is, and what impact it has. That is, neighborhood context is tautologically defined (partly) by the characteristics and behaviors of its residents. Neighborhood context can, however, shape these behaviors (both through social interactions transpiring internally and exogenous forces) and, thus, some characteristics of its residents. Neighborhood context also will influence how many and which current residents will move out (both by altering residents' characteristics and their satisfaction with the context) and how many residents with what characteristics will move in to replace them. How much residential turnover and how many alterations of the socioeconomic and demographic character of a neighborhood occur will, in turn, affect the nature and stability of the neighborhood context over time and its ability to exert independent behavioral and other effects on its residents. Having introduced this nexus of causal interconnections, we now turn to a review of the empirical evidence produced on both sides of the Atlantic that have contributed to our understanding of these relationships.

## Neighborhood Effects and Connections to Residential Mobility

Residential segregation and resulting neighborhood concentrations of households distinguished by their racial, ethnic, or socioeconomic characteristics have been of academic interest for generations. The causes and characteristics of segregation have long claimed the most attention, but much recent research has also been devoted to understanding consequences in terms of neighborhood effects. In an often-cited article, Galster and Killen (1995) argue that neighborhoods affect the opportunities and life chances of their inhabitants through various mechanisms like provision of services, effects on attitudes of others, and endogenous processes affecting individual decisionmaking, preferences, and perceived opportunities. Following this argument, residential context at various spatial scales delineates an "opportunity structure" whereby some neighborhoods affect their inhabitants positively while others restrain opportunities. Since Wilson's 1987 publication, much focus has been directed to concentrated poverty areas as these were alleged to restrict residents' opportunities the most and produce the most negative externality effects for the larger society.

The number of empirical studies testing the neighborhood effect hypothesis has rapidly increased since the early 1990s, testing outcomes like income (Oreopoulos 2003; Galster, Marcotte, et al. 2007; Galster et al. 2008; Galster, Andersson, and Musterd 2010), employment (Musterd, Ostendorf, and de Vos 2003; Dawkins, Shen, and Sanchez 2005; Musterd and Andersson 2006; van Ham and Manley 2010), educational attainment (Crane 1991; Harding 2003; Crowder and South 2011; Wodtke, Harding, and Elwert 2011; Sykes and Musterd 2011), welfare usage (Vartanian and Gleason 1999), crime (Kling, Ludwig, and Katz 2005), teenage sexuality and parenthood (Crane 1991; Brooks-Gunn et al. 1993), health (Propper et al. 2005; Ross and Mirowsky 2008), cognitive development (Sampson, Sharkey, and Raudenbush 2008), and acceptance of "deviant behavior" (Friedrichs and Blasius 2003). A few studies have also looked at how neighborhood context affects individual behaviors shaping the physical environment of the neighborhood (Grigsby et al. 1987; Taub, Taylor, and Dunham 1984). For example, Galster (1987) demonstrated how the home maintenance behavior of owner–occupiers is influenced by their neighborhood social interactions and attachments. This statistical literature on neighborhood effects has been summarized and evaluated by several reviews (Jencks and Mayer 1990; Friedrichs 1998; Leventhal and Brooks-Gunn 2000; Sampson, Morenoff, and Gannon-Rowley 2002; Friedrichs, Galster, and Musterd 2003; Galster 2008; DeLuca and Dayton 2009), and although findings are often divergent, some consensual conclusions can be drawn. Most importantly, most US studies identify neighborhood effects.[3] They typically suggest that poverty areas have negative effects on a variety of outcomes for both children and adults, although the effects on children are generally more pronounced. The European literature is less consistent, and effects found are typically smaller. Results are also affected by the scale on which neighborhoods are defined, the characteristics of the population being affected (demographic, socioeconomic, ethnic and racial), and the macro context in which the neighborhood is embedded (Galster et al. 2010).

Thresholds and nonlinear correlations between a neighborhood socioeconomic feature and various outcomes are also increasingly common in the neighborhood effects literature. In his review of the U.S. literature, Galster (2002) finds consistent threshold effects on various resident behaviors when the neighborhood poverty rates exceed 15 to 20 percent. Galster et al. (2010) and Hedman and Galster (2013) find substantial negative effects on individual Stockholm residents' incomes if they reside in neighborhoods with over 40 percent low-income neighbors.

Often embedded in neighborhood effect theory is the implicit assumption that neighborhoods are static. Neighborhoods are often discussed as separate entities providing a relatively constant context that consistently affects individuals who are exposed to it for an extended period (e.g., childhood upbringing). The neighborhood's status in the urban hierarchy is typically regarded as fixed over time, as are the features that make up its opportunity structures. For example, an area once defined as a poverty area is implicitly assumed to remain a poverty area, with essentially unaltered or only slowly changing physical and population characteristics.

Of course, neighborhoods typically are dynamic; they often change their aggregate demographic and socioeconomic profiles with frequently selective moves in and out by households. The selectivity of who stays, who moves in, and who moves out can maintain, improve, or impair a neighborhood's status position in the urban hierarchy and its internal social dynamics. But, more fundamentally, the dynamism of the neighborhood context can directly shape the magnitude of any measured neighborhood effect by shaping the duration of exposure (Galster 2012).

The temporal dimension of neighborhood effects has been made more explicit in limited empirical work, with a few studies paying attention to how variations in the timing and duration of exposure modified the observed relationship. They paint a consistent portrait, however, that neighborhood effects seem to be stronger if the exposure is cumulative, and sometimes effects appear only after a lag. Aaronson (1998) examined how neighborhood poverty rates affected teens' school dropout rates and found that the average (cumulative) neighborhood conditions experienced between ages 10 and 18 were much stronger predictors than contemporaneous conditions. Wheaton and Clarke (2003) investigated the temporal dimension of neighborhood disadvantage effects on the mental health of young adults. They found that current neighborhood had no effect, but earlier neighborhood disadvantage experienced as a child had a lagged effect that grew stronger as cumulative exposure intensified. Turley (2003) found that white (though not black) children's school test scores and several behavioral indicators grew more efficacious the greater the mean income of their neighborhoods. These relationships were strongest for children who had lived in their neighborhoods for three years or more, suggesting either a lagged and/or cumulative effect process. Kauppinen (2007) observed little impact of neighbors' social status on type of secondary school chosen unless the students were in the neighborhood two or more years. Sampson et al. (2008) examined reading abilities of black children who grew up in Chicago at three later points in their lives. Their findings indicated that there was a cumulative, durable penalty from extended childhood residence in neighborhoods with concentrations of low socioeconomic–status households that grew stronger after several years of residence in such places. Finally, Musterd, Galster, and Andersson (2012) investigated the effect of neighborhood income mix on individuals' earnings. They found important temporal dimensions in the statistical effect of neighborhood income mix: Recent, continued, or cumulative exposure yielded stronger associations than lagged, temporary ones, and there was a distinct time decay (though some persistence) in the potential effects after exposure ceased (though with some gender differences).

Finally, we want to raise the possibility that residential mobility itself could be affected by neighborhood effect mechanisms in the same way as other behaviors or outcomes. Following neighborhood effect theory, there is no reason to believe that externality effects on preferences, aspirations, norms and values, and subjective perceptions of possible outcomes should not include preferences and perceptions about individuals' own and other neighborhoods, as well as about mobility in general, nor that any opportunity structure should not affect opportunities in terms of if and where to move. Exogenous and correlated effects are already tested to some extent, although studies have not made this explicit claim. The white flight theory describes how exogenous characteristics of others affect moves; subjective neighborhood evaluations can be seen as a result of public services, location, and neighborhood characteristics associated with correlated mechanisms; and neighborhood reputation and stigmatization are closely related. Vartanian, Walker Buck, and Gleason (2007) have also tested place socialization theory when showing that children growing up in poverty areas are more likely than others to live in such areas as adults, all else being equal.

To sum up this section, considerable research shows that several aspects of neighborhood context can produce nontrivial independent effects on a variety of outcomes for resident children, youth, and adults, although the impacts may be observed only after certain threshold values of context have been exceeded. However, the context is intimately connected with residential mobility, as we explore in the next section.

## Residential Mobility in Relation to Neighborhoods

The dynamics of neighborhoods embody two different but related themes: the neighborhood's status and characteristics compared with other neighborhoods (city dynamics) and dynamic processes within the specific neighborhood. Residential mobility is a key to both. Mobility is studied both at the micro and macro level. The former is primarily concerned with preferences and constraints of individual movers, and the latter studies how these preferences and constraints are transformed into urban moving patterns and flows in the aggregate, often in relation to residential segregation by ethnicity or income. Selective mobility patterns on the macro level and self-selection into neighborhoods on the micro level constitute keys to understanding how residential segregation and neighborhood inequality are maintained or altered over time. Although residential segregation encompasses several dimensions, studies have most often focused on the racial and ethnic dimension: the production and reproduction of ethnic neighborhoods[4] through ethnically selective moves.

Traditionally, explanations of ethnic segregation focused on the moves of minorities, claiming that segregation was a result of voluntary ethnic clustering due to preferences for living near coethnics and the putative benefits gained from such a living. However, these theories have received much criticism for being built on racist ideology and having little empirical support (Smith 1989; Molina 1997). More common explanations today point at structural factors constraining the choice sets of minorities and low-income people [see Bolt and van Kempen (2003) for an overview]. A complementary body of literature looks at other forms of constraints, especially ethnic discrimination, which has been confirmed in several countries; see, for example, Ross and Turner (2005) for the United States and Robinson (2002) for the United Kingdom. Yet another body of research focuses on interethnic differences in socioeconomic status, arguing that minorities' lower socioeconomic status explains their overrepresentation in certain areas (Alba and Logan 1993). The emphasis on economic and discriminatory constraints faced by ethnic minorities has led to a shift in focus toward the moves of those experiencing fewer constraints on the housing market: the ethnic majority and better-off households.

In the United States, the white flight theory has become the most influential in explaining ethnic residential segregation. The theory explains changes in ethnic composition through the "flight" of white inhabitants when the share of black [or other minorities: see Clark (1992) and Pais, South, and Crowder (2009)] residents reaches a critical tipping point (Galster 1991; Clark 1992; South and Crowder 1997, 1998; Crowder 2000; Ellen 2000; Quillian 2002; Card, Mas, and Rothstein 2008). Proposed reasons for this out-mobility of the white population are racism (Farley et al. 1994) and a fear of dropping housing prices (Harris 1999). White avoidance, a complementary theory explaining how ethnic minority areas are reproduced (rather than come into being), emphasizes white people's unwillingness to enter neighborhoods with a certain share of blacks or other minorities (Ellen 2000).[5] Similar discussions in the literature focus on how poverty areas either come into being or are made even poorer through selective moving patterns in which better-off residents leave and are replaced by households with a socioeconomic status similar to nonmovers (Wilson 1987; Friedrichs 1991; Skifter Andersen 2003; Andersson and Bråmå 2004). A high turnover rate might speed up this process and may also be perceived as a sign of low attractiveness, thereby reinforcing the selective mobility patterns (Andersson and Bråmå 2004). The process can of course be reversed, as in the gentrification example; neighborhood investments in home rehabilitations or new construction may cause housing prices to rise, forcing low-income groups to leave and be replaced by better-off inhabitants.

The above theories of ethnic and socioeconomic selective mobility suggest that moving patterns are affected by thresholds at which the share of ethnic minority or low-income inhabitants affects aggregate moving patterns as it exceeds a critical value [for a review, see Quercia and Galster (2000)]. Though accepted as conventional wisdom for a long prior period (Wolf 1963), a theoretical foundation for nonlinear change processes in a neighborhood's racial composition did not emerge until Schelling's tipping model (1971), which was subsequently extended by Taub et al. (1984). Numerous empirical studies have indeed found that threshold-like relationships characterize neighborhood racial transitions, though no universal tipping point exists (Giles 1975; Goering 1978; Galster 1991; Clark 1991; Lee and Wood 1991; Crowder 2000; Card et al. 2008). Indeed, an even wider range of neighborhood sociodemographic dynamics may be associated with nonlinear processes (Galster, Quercia, and Cortes 2000; Galster, Cutsinger, and Lim 2007; Lim and Galster 2009).

To understand the origins of the aggregate flows of majority (and minority) households leading to selective moving patterns and residential segregation, one must look at the micro level and mobility decisions of individual households. Much literature on the micro level follows Brown and Moore's (1970) classic work and divides the moving process into two separate decisions: the decision to move and the choice of destination. These decisions are of course interrelated as many people decide to move when they find an interesting and feasible alternative. Nevertheless, the distinction is useful for theoretical purposes.

The decision to move is often viewed on the micro level as an outcome of a primary dissatisfaction with the home in relation to the households' preferences and needs (Brown and Moore 1970; Speare 1974; Clark and Dieleman 1996). A growing body of literature suggests, however, that more focus should be directed to the role of the neighborhood (Lee, Oropresa, and Kanan 1994; Lu 1998; Kearns and Parkes 2003; Clark, Deurloo, and Dieleman 2006; Clark and Ledwith 2006; van Ham and Feijten 2008; Feijten and van Ham 2009; van Ham and Clark 2009; Hedman, van Ham, and Manley 2011). That the neighborhood matters for residential satisfaction and mobility decisions has been shown empirically. When Coleman (1978) asked people whether they preferred a good house in a less desirable neighborhood or a less desirable house in a good neighborhood, a vast majority chose the latter alternative. Similarly, results by Galster (1987), Lee et al. (1994), and Lu (1998) showed that people who were satisfied with their neighborhood environment were less likely to report moving intentions than those who were dissatisfied. The factors affecting residential satisfaction are of course many and likely to differ among individuals and groups depending on preferences and what they are used to (Galster and Hesser 1981). Several studies, however, have shown a positive correlation between neighborhood attachment and neighborhood satisfaction and between neighborhood attachment, moving intentions, and actual mobility (Clark and Ledwith 2006; Guest et al. 2006; Permentier, van Ham, and Bolt 2009). People who are attached to their neighborhoods, either to the area itself or to people or activities in it, are less likely to leave. Residential dissatisfaction seems to be greatest in poverty areas (Burrows and Rhodes 2000; Parkes, Kearns, and Atkinson 2002; Kearns and Parkes 2003). This finding can probably be related to the generally high impact of physical appearance, quality and number of services, and crime rates on residential satisfaction. Permentier et al. (2007, 2009) also point specifically to the reputation of neighborhoods, showing that perceived neighborhood reputation is a significant predictor of moving intentions (Permentier et al. 2009).

There is also a literature suggesting that moving plans are not only caused by neighborhood conditions per se but also by neighborhood change (Taub et al. 1984; Galster 1987; Lee et al. 1994; Kearns and Parkes 2003; Feijten and van Ham 2009; van Ham and Clark 2009). The white flight theory suggesting that changes in ethnic composition affect mobility behavior is supported by three recent micro-level studies on Dutch data by van Ham and Feijten (2008), Feijten and van Ham (2009), and van Ham and Clark (2009). They all found a positive relationship between moving wishes and an increase in the percentage of ethnic minorities, but this correlation was much less pronounced for people belonging to an ethnic minority themselves. The results by Feijten and van Ham (2009) showed correlations between moving intentions and a decrease in the neighborhood's socioeconomic status (based on mean income, mean education level, and mean unemployment rate); an increase in status reduced wishes to leave.

The choice of destination, the second step in Brown and Moore's model, can in an intraurban perspective be seen in basically the same residential (dis)satisfaction framework.[6] Alternatives are evaluated against factors important to the household, the ultimate destination being the one thought to provide the most prospective satisfaction among the alternatives open to the mover. However, as noted, the number of potential alternatives varies widely among households due to constraints that can be financial, related to the housing market (stock, composition, turnover rate) or housing allowance systems (housing queues, other distribution mechanisms), or be discriminatory in character and thus especially limit the choice set for ethnic minorities and low-income people. Several empirical studies from both sides of the Atlantic have found that these groups react to declining neighborhood conditions in the same way as other groups but are less likely to transform this dissatisfaction into a move (South and Crowder 1997, 1998; Bolt and van Kempen 2003; Kearns and Parkes 2003) or, if moving, move between areas of similar characteristics (Bråmå and Andersson 2005). Constraints can also be thought of in terms of information, stereotypes, tradition, and other factors directing the household to some alternatives while excluding others. The literature on long-distance migration has shown how migrants tend to follow the paths of previous migrants[7] as a result of information spread within networks, facilitating mechanisms, assistance, and acceptance of both migration behavior and the destination (Cadwallader 1992). Although not directly applicable to the intraurban scale, factors such as information, social and family networks, and acceptance of a destination among peers are still important for destination choice (Hedman 2013).

Finally, we must note that aggregate mobility rates per se can affect the neighborhood, independent of any selective migration that may alter the neighborhood's composition. For example, Sampson and colleagues argue that high mobility rates disrupt social organization and collective efficacy, thereby affecting community norms and values. They have found a positive correlation between high turnover rates and criminality levels in neighborhoods (Sampson and Groves 1989; Sampson et al. 2002).

To sum up this section, research has shown that several dimensions of the neighborhood affect moving decisions and destination choices of households, but the ability to put preferences into practice varies because constraints are more severe for some groups. On the aggregate level, preferences and constraints produce selective moving patterns and ultimately produce residential segregation involving neighborhoods of distinctive demographic characteristics and socioeconomic status. Continued selective mobility tends to reproduce segregation patterns, but that result is not a given, as shown by studies of white flight and neighborhood change, neither does it mean that neighborhoods are static. The profiles of neighborhoods can sometimes change quickly when out-migration tipping points are exceeded. And even when their profiles are not altered, there can be high rates of turnover in even the poorest neighborhoods (Quillian 2002; Andersson 2008). These dynamic places form the contexts where neighborhood effects may take place, but extremely high rates of resident mobility may nullify any effect of context if they either render individual durations of exposure times insufficient or produce an inconsistent context to which residents are exposed.

## Research Challenges from the Perspective of the Holistic Model

Thus far we have argued that sufficient empirical evidence exists to suggest that our holistic approach to neighborhood change, neighborhood effects, and residential mobility (as embodied in equations 1 to 7) can be taken as a working hypothesis. If so, what are the implications for quantitative research in these realms? There are many (Galster 2008; Hedman 2011), but here we focus on two salient challenges related to correcting potential biases. The first challenge is selective mobility into and out of neighborhoods based on unobserved characteristics of households. The second challenge is endogeneity: neighborhood context and residential mobility are mutually causal. The next two subsections discuss these challenges in more detail and how the empirical literature has dealt with them.

## Challenge 1: Geographic Selection Bias from Unobservables

The first challenge is geographic self-selection: moves from and to places are not made randomly. Instead, households choose to stay or to leave their neighborhood and then select another neighborhood for at least some important reasons that often are unknown to the researcher. If the reasons behind this choice are omitted from the analysis and also are associated with the individual outcome under investigation, results will be biased[8] (Jencks and Mayer 1990; Tienda 1991; Manski 2000; Duncan, Connell, and Klebanov 1997; Galster 2008). For example, parents with high ambitions for their child likely will choose a neighborhood with "good" potential peers. If the grades of the child are not only affected by the neighborhood peers but also by the ambitious parents (a factor not controlled for), correlations attempting to measure the causal role of the former will be biased.

This challenge is well-known in the neighborhood effects literature. Most recent neighborhood effects studies based on nonexperimentally produced data control for selection bias through a variety of econometric techniques; for more on the alternative strategies, see Galster (2008). Although several of these studies still find evidence of neighborhood effects after geographic selection is controlled (Dawkins et al. 2005; Galster, Marcotte et al. 2007; Galster et al. 2008; Galster, Andersson and Musterd, 2010; Musterd, Andersson and Galster 2012), others argue that selection effects dominate true causal effects (Oreopoulos 2003; van Ham and Manley 2010). All these econometric strategies have their strengths and weaknesses, a full discussion of which is beyond the scope of this essay.

Experimental designs or "natural quasi-experiments" have also been used to overcome selection bias. In random assignment experiments, data are produced by a design whereby households are randomly assigned to different neighborhoods, such as the Moving to Opportunity demonstration (Ludwig, Ladd, and Duncan 2001; Orr et al. 2003; Kling, Liebman, and Katz 2007; Ludwig et al. 2008). In natural experiments, observations are produced by idiosyncratic interventions, such as the Gautreaux and Yonkers public housing desegregation programs (Rosenbaum 1995; DeLuca et al. 2010), public housing revitalization programs (Jacob 2004; Clampet-Lundquist 2007), or inclusionary zoning (Schwartz 2010) that creates exogenous variation in neighborhood environments for assisted tenants. These strategies for testing for neighborhood effects have also reached different conclusions, depending on the outcome of interest. Not surprisingly, they also have their respective strengths and weaknesses (Galster 2008, 2011).

The above methods arguably serve their purpose in reducing selection bias, but they do not tell us anything about selection processes, which should be of interest to neighborhood effects scholars [see also Sampson and Sharkey (2008)]. Few neighborhood effects studies have tried to control for selection by attempting to model destination choices of individuals. One important exception is Ioannides and Zabel's (2008) study on how neighborhood externalities affect housing structure demand. Ioannides and Zabel use a two-step model in which the first step models neighborhood choice and the second uses these estimates as a control variable[9] for selection bias when analyzing outcomes of housing demand. Their results showed that households move to places inhabited by people similar to themselves in terms of income, age, race, and so on (i.e., those variables controlled for in the model). However, the explanatory power of the model was not very high (0.16 when using interaction variables), indicating that other important sorting mechanisms were not included in the study. Regardless, the results in the second stage of their analysis showed that controlling for selection had large impacts on results, in this case strengthening the apparent effects of neighborhood characteristics on housing structure demand. Although not yet perfect, their two-step model is a promising strategy as it not only controls for selection bias when measuring neighborhood effects, but also sheds more light on how this selection process works.

If the neighborhood effects literature can be criticized for devoting too little effort to consider the explicit modeling of selection, the residential mobility literature can be more severely criticized for overlooking entirely the potential bias from geographic selection based on unobserved characteristics of households. In any statistical study of residential mobility there likely will be a set of unobserved individual or household characteristics that influences both selection into and out of neighborhoods and observed individual behaviors. In these circumstances, partial correlations between neighborhood characteristics and individual characteristics will not provide unbiased estimates of the true magnitude of the causal influence of one on the other, regardless of which direction of causation is posited. An illustration of an unobserved characteristic is the salience given to visible symbols of prestige; those placing great weight on such symbols will work harder to evince higher incomes and will try to live in prestigious neighborhoods. In a model testing the effect of individuals' incomes on the prestige of their neighborhoods, how much causal impact can be rightfully attributed to the former with "prestige salience" uncontrolled?

## Challenge 2: Endogeneity

The endogeneity problem refers to the mutual causality of individual (household) characteristics and associated neighborhood characteristics.[10] As is apparent from the italicized terms in equations 1 to 7, neighborhood changes, neighborhood effects on residents, and household mobility patterns both affect and are affected by each other. In statistical terms, this means that error terms are correlated among the various equations, which produces biased estimates of the coefficients of the endogenous variables. We emphasize that this likely source of bias plagues both the neighborhood effects and the residential mobility literatures, though it has almost never been recognized (let alone confronted statistically) in either (Hedman 2011; Hedman and Galster 2013).

## Meeting the Dual Challenges of Geographic Selection and Endogeneity Biases

Although the neighborhood effects literature has devoted substantial effort to reducing geographic selection bias through econometric techniques, natural experiments, and (in one case) a random assignment experiment, neither it nor the residential mobility literature has taken the threat of endogeneity bias seriously. We think that the cutting edge of both fields lies in the holistic econometric modeling of relationships such as those sketched in equations 1 to 7. The classic method for dealing with multiple equations embodying endogenous variables is to employ instrumental variables (IVs) in place of the endogenous predictors. There is an added benefit by doing so here: IV approaches also provide a means of subduing bias from geographic selection on unobservables (Galster, Marcotte et al. 2007). Three prototype efforts are promising.

Galster, Marcotte, et al. (2007) developed a model in which parental housing tenure, expected length of stay, and neighborhood poverty rate were endogenous over the first 18 years of a child's lifetime and, in turn, jointly affected their outcomes measured as young adults. They used two-stage least squares to obtain IV estimates for shares of childhood years spent in a home owned by parents, shares of childhood years when there were no residential moves, and mean neighborhood poverty rate experienced during childhood. These IV estimates were employed in second-stage equations predicting education, fertility, and young-adulthood income outcomes. The results indicated that being raised in higher-poverty neighborhoods had a substantial negative effect on educational attainments and indirectly on incomes. Though emphasizing neighborhood effects, this holistic approach had the potential of also providing insights about the mobility behavior of households, though these IV results were not reported.

Sari (2012) developed a two-equation simultaneous model using 1999 data on Paris-region adult males in which residence or nonresidence in a deprived neighborhood was one outcome, being employed or unemployed was the other, and the two were mutually causal. Sari employed a bivariate probit maximum likelihood method with an IV to obtain an unbiased estimate of the effect of deprived neighborhood residence on an individual's employment probability. He found that residence in a deprived neighborhood was associated with substantially lower employment probabilities. Unfortunately, Sari did not attempt to estimate the deprived neighborhood residence equation with an IV for employment, so a holistic empirical portrait is missing.

Most recently, we (Hedman and Galster 2013) estimated a structural equation system in which neighborhood income mix and individual resident income were specified as mutually causal. Statistical tests using data from Stockholm verified the endogenous nature of these predictors. By adopting a fixed-effects panel model with IV proxies of both endogenous predictors, we addressed both sources of bias. We found that selection both on unobservables and endogeneity were empirically important sources of potential bias in studies of neighborhood effects and neighborhood selection. Our IV approach produced a substantially different pattern of parameter estimates than a conventional fixed-effect model. First, coefficients were larger (with the exception of income predicting low-income share), implying that endogeneity biased measured neighborhood effect and neighborhood income sorting processes downwards. Second, our IV model found powerful evidence of nonlinear patterns not apparent in the fixed-effect specification. There was no statistically significant negative effect on individual income from the percentage of low-income neighbors before the share reached 20 percent; there was another threshold at 50 percent low-income neighbors at which the negative effects became much larger. We also found that neighborhood income sorting becomes intensified when individual income is greater than the mean, with not only marginal increases in high-income neighbors but marginal decreases in low-income neighbors. By implication, our study suggests that selection and endogeneity biases can be severe enough to substantially distort findings.

We of course recognize that the use of IVs raises its own particular challenges. To be an appropriate and efficacious technique, IVs must be both valid and strong (Murray 2006). To be valid, IVs must (1) be uncorrelated with the error term, (2) be correlated with the endogenous variable, and (3) not be otherwise included in the given equation. The strength of the IVs is the degree to which they are correlated with the endogenous variable for which they stand in. The central challenge of any empirical effort is to identify and measure variables that can serve as IVs and meet these criteria. Though in our recent work (Hedman and Galster 2013) our IVs passed most of the standard tests for strength and validity, the method always should be subject to some skepticism. Nevertheless, we think it holds promise for simultaneously dealing with the twin challenges of selection on unobservables and endogeneity.

## Conclusion

In this essay we argue that neighborhood conditions and their dynamics, neighborhood effects on residents, and residential mobility are intrinsically related through numerous causal connections. We offer a set of seven stylized equations that illustrate this point and demonstrate that there is strong support for this view in the scholarly literatures of the United States and Western Europe. Although we suspect that our formulation of these interconnections is not controversial, its implication for empirical studies perhaps is more so. We show that statistical studies in both the fields of neighborhood effects and residential mobility potentially suffer from biasing forces related to geographic selection on unobserved individual characteristics and endogeneity. Though the former field has devoted a good deal of effort to develop methods to overcome selection bias, the latter field has almost universally overlooked it. Neither field has made much effort to address the challenge of endogeneity.

We believe that prototype efforts to model neighborhood effects and mobility holistically through the use of IVs represent the cutting edge of the next generation of studies in both fields. We have offered our own work as an example of the important empirical difference that such an approach produces, demonstrating that our concerns over bias are not just theoretical or hypothetical. We recognize that the IV approach is rife with challenges of its own. Paramount is the difficulty in identifying valid and powerful variables to serve as instruments. Nevertheless, we believe that the potential payoffs from further explorations in this realm should prove worthwhile.

## NOTES

1. The need for a more holistic and dynamic approach to neighborhoods has previously been emphasized by scholars interested in neighborhood effects (Galster 1987, 2003; Tienda 1991; Galster, Marcotte, et al. 2007). Ioannides and Zabel (2008), van Ham and Manley (2010), Hedman and van Ham (2011), and Hedman, van Ham, and Manley (2012) have argued that one must uncover neighborhood selection processes to accurately assess neighborhood effects. They do not, however, holistically extend the argument conceptually or empirically to consider how neighborhood effects may alter neighborhood selection.

2. We recognize that perceptions of alternative neighborhoods are based on imperfect information and that there are spatial biases to this information. A fuller explanation of the residential search process is beyond the scope of this essay.

3. Despite its vaunted reputation, the Moving to Opportunity demonstration research has major flaws that render any conclusions regarding neighborhood effects based on its data highly suspect. For a thorough review see Galster (2011).

4. We here refer both to neighborhoods with high shares of ethnic minorities and neighborhoods inhabited by an ethnically homogenous majority (native) population.

5. See Bråmå (2006) for an attempt to apply the white flight and white avoidance theories to a Swedish context.

6. The framework is more problematic for long-distance migration (or moves within very large metropolitan areas) as other factors come into play; see Roseman (1983).

7. This is referred to as chain migration or migration network theory.

8. The direction of these potential biases is, however, unclear; Jencks and Mayer (1990) and Tienda (1991) argue that effects are overestimated, but Brooks-Gunn, Duncan, and Aber (1997) suggest that the opposite could also occur.

9. A "Heckman correction" variable; see various papers on selection bias by James Heckman from 1979 onward.

10. There is another sort of endogeneity that we do not discuss here. The choice of neighborhood is undertaken as a joint decision including decisions about housing, tenure, and length of residence (Galster 2003, 2008).

## REFERENCES

Aaronson, D. 1998. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *Journal of Human Resources* 33: 915–46.

Alba, R. D., and J. R. Logan. 1993. "Minority Proximity to Whites in Suburbs: An Individual Level Analysis of Segregation." *American Journal of Sociology* 98: 1388–1427.

Andersson, R. 2008. "Neighbourhood Effects and the Welfare State: Towards a European Research Agenda?" *Schmollers Jahrbuch* 128: 1–14.

Andersson, R., and Å. Bråmå. 2004. "Selective Migration in Swedish Distressed Neighbourhoods: Can Area-Based Urban Policies Counteract Segregation Processes?" *Housing Studies* 19 (4): 517–39.

Bolt, G., and R. van Kempen. 2003. "Escaping Poverty Neighbourhoods in the Netherlands." *Housing, Theory and Society* 20: 209–22.

Bråmå, Å. 2006. "'White Flight'? The Production and Reproduction of Immigrant Concentration Areas in Swedish Cities 1990–2000." *Urban Studies* 43 (7): 1127–46.

Bråmå, Å., and R. Andersson. 2005. "Who Leaves Sweden's Large Housing Estates?" In *Restructuring Large Housing Estates in Europe,* edited by R. van Kempen, K. Dekker, S. Hall, and I. Tosics. Bristol, United Kingdom: Policy Press.

Brooks-Gunn, J., G. Duncan, and J. Aber. 1997. *Neighborhood Poverty: Volume 1, Context and Consequences for Children.* New York: Russell Sage Foundation.

Brooks-Gunn, J., G. Duncan, P. Kato Klebanov, and N. Sealand. 1993. "Do Neighborhoods Influence Child and Adolescent Development?" *American Journal of Sociology* 99 (2): 353–95.

Brown, L. A., and E. G. Moore. 1970. "The Intra-Urban Migration Process: A Perspective." *Geografiska Annaler B* 52 (1): 1–13.

Burrows, R., and D. Rhodes. 2000. "The Geography of Misery: Area Disadvantage and Patterns of Neighbourhood Dissatisfaction in England." In *Researching Poverty,* edited by J. Bradshaw and R. Sainsbury. Aldershot, United Kingdom: Ashgate.

Cadwallader, M. 1992. *Migration and Residential Mobility: Macro and Micro Approaches.* Madison, WI: University of Wisconsin Press.

Card, D., A. Mas, and J. Rothstein. 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics* 123 (1): 177–218.

Clampet-Lundquist, S. 2007. "No More 'Bois Ball: The Impact of Relocation from Public Housing on Adolescents." *Journal of Adolescent Research* 22 (3): 298–323.

Clark, W.A.V. 1991. "Residential Preferences and Neighborhood Racial Segregation: A Test of the Schelling Segregation Model." *Demography* 28 (1): 1–19.

Clark, W.A.V. 1992. "Residential Preferences and Residential Choices in a Multiethnic Context." *Demography* 29: 451–66.

Clark, W.A.V., M. Deurloo, and F. Dieleman. 2006. "Residential Mobility and Neighbourhood Outcomes." *Housing Studies* 21 (3): 232–342.

Clark, W.A.V., and F.M. Dieleman. 1996. *Households and Housing: Choice and Outcomes in the Housing Market.* New Brunswick, NJ: CUPR Press.

Clark, W.A.V., and V. Ledwith. 2006. "Mobility, Housing Stress and Neighbourhood Contexts: Evidence from Los Angeles." *Environment and Planning A* 38: 1077–93.

Coleman, R. 1978. "Attitudes toward Neighborhoods: How Americans Choose to Live." Working paper no. 49. Cambridge, MA: Joint Center for Urban Studies, Massachusetts Institute of Technology, and Harvard University.

Crane, J. 1991. "The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing." *American Journal of Sociology* 96 (5): 1226–59.

Crowder, K. 2000. "The Racial Context of White Mobility: An Individual-Level Assessment of the White Flight Hypothesis." *Social Science Research* 29: 223–57.

Crowder, K., and S.J. South. 2011. "Spatial and Temporal Dimensions of Neighborhood Effects on High School Graduation." *Social Science Research* 40: 87–106.

Dawkins, C.J., Q. Shen, and T.W. Sanchez. 2005. "Race, Space, and Unemployment Duration." *Journal of Urban Economics* 58: 91–113.

DeLuca, S., and E. Dayton. 2009. "Switching Contexts: The Effects of Housing Mobility and School Choice Programs on Youth Outcomes." *Annual Review of Sociology* 35: 457–91.

DeLuca, S., G. Duncan, R. Mendenhall, and M. Keels. 2010. "Gautreaux Mothers and Their Children: An Update." *Housing Policy Debate* 20 (1): 7–25.

Doff, W. 2010. *Puzzling Neighbourhood Effects.* Amsterdam, Netherlands: IOS Press.

Duncan, G., J. Connell, and P. Klebanov. 1997. "Conceptual and Methodological Issues in Estimating Causal Effects of Neighborhoods and Family Conditions on Individual Development." In *Neighborhood Poverty: Volume 1, Context and Consequences for Children,* edited by J. Brooks-Gunn, G. Duncan, and J. Aber. New York: Russell Sage Foundation.

Ellen, I. 2000. *Sharing America's Neighborhoods.* Cambridge, MA: Harvard University Press.

Farley, R., C. Steeh, M. Krysan, T. Jackson, and K. Reeves. 1994. "Stereotypes and Segregation: Neighbourhood in the Detroit Area." *American Sociological Review* 44: 425–28.

Feijten, P., and M. van Ham. 2009. "Neighbourhood Change ... Reason to Leave?" *Urban Studies* 46 (10): 2103–22.

Friedrichs, J. 1991. "Middle-Class Leakage in Large New Housing Estates: Empirical Findings and Policy Implications." *Journal of Architectural and Planning Research* 8: 287–95.

Friedrichs, J. 1998. "Do Poor Neighbourhoods Make Their Residents Poorer? Context Effects of Poverty Neighbourhoods on Residents." In *Empirical Poverty Research* *in Comparative Perspective,* edited by H. J. Andress. Aldershot, United Kingdom: Ashgate.

Friedrichs, J., and J. Blasius. 2003. "Social Norms in Distressed Neighbourhoods: Testing the Wilson Hypothesis." *Housing Studies* 18 (6): 807–26.

Friedrichs, J., G. Galster, and S. Musterd. 2003. "Neighbourhood Effects on Social Opportunities: The European and American Research and Policy Context." *Housing Studies* 18 (6): 797–806.

Galster, G. 1987. *Homeowners and Neighborhood Reinvestment.* Durham, NC: Duke University Press.

Galster, G. 1991. "White Flight from Integrated Neighborhoods." *Urban Studies* 27 (3): 385–99.

Galster, G. 2001. "On the Nature of Neighborhood." *Urban Studies* 38 (12): 2111–24.

Galster, G. 2002. "An Economic Efficiency Analysis of Deconcentrating Poverty Populations." *Journal of Housing Economics* 11 (4): 303–29.

Galster, G. 2003. "Investigating Behavioural Impacts of Poor Neighbourhoods: towards New Data and Analytic Strategies." *Housing Studies* 18 (6): 893–914.

Galster, G. 2008. "Quantifying the Effect of Neighbourhood on Individuals: Challenges, Alternative Approaches and Promising Directions." *Journal of Applied Social Sciences. Schmollers Jahrbuch* 128 (1): 7–48.

Galster, G. 2011. "Changing the Geography of Opportunity by Helping Poor Households Move Out of Concentrated Poverty: Neighborhood Effects and Policy Design." In *Neighborhood and Life Chances: How Place Matters in Modern America,* edited by H. Newburger, E. Birch, and S. Wachter (221–34). Philadelphia, PA: University of Pennsylvania Press.

Galster, G. 2012. "The Mechanisms of Neighbourhood Effects: Theory, Evidence, and Policy Implications." In *Neighbourhood Effects Research: New Perspectives,* edited by M. van Ham, D. Manley, N. Bailey, L. Simpson, and D. Maclennan (23–56). Dordrecht, Netherlands: Springer.

Galster, G., R. Andersson, and S. Musterd. 2010. "Who Is Affected by Neighborhood Income Mix? Gender, Age, Family, Employment and Income Differences." *Urban Studies* 47 (14): 2915–44.

Galster, G., R. Andersson, S. Musterd, and T. M. Kauppinen. 2008. "Does Neighborhood Income Mix Affect Earnings of Adults? New Evidence from Sweden." *Journal of Urban Economics* 63: 858–70.

Galster, G., J. Cutsinger, and U. Lim. 2007. "Are Neighborhoods Self-Stabilizing? Exploring Endogenous Dynamics." *Urban Studies* 44 (1): 1–19.

Galster, G., and G. Hesser. 1981. "Residential Satisfaction: Contextual and Compositional Correlates." *Environment and Behavior* 13: 735–58.

Galster, G., and S. Killen. 1995. "The Geography of Metropolitan Opportunity: A Reconnaissance and Conceptual Framework." *Housing Policy Debate* 6 (1): 7–43.

Galster, G., D. E. Marcotte, M. Mandell, H. Wolman, and N. Augustine. 2007. "The Influence of Neighborhood Poverty during Childhood on Fertility, Education, and Earnings Outcomes." *Housing Studies* 22 (5): 723–51.

Galster, G., R. Quercia, and A. Cortes. 2000. "Identifying Neighborhood Thresholds: An Empirical Exploration." *Housing Policy Debate* 11 (3): 701–32.

Giles, M. 1975. "White Flight and Percent Black: The Tipping Point Re-examined." *Social Science Quarterly* 56 (1): 85–92.

Goering, J. 1978. "Neighborhood Tipping and Racial Transition: A Review of Social Science Evidence." *Journal of the American Institute of Planners* 44: 68–78.

Grigsby, W., M. Baratz, G. Galster, and D. Maclennan. 1987. *The Dynamics of Neighborhood Change and Decline.* London: Pergamon Press.

Guest, A. M., J. K. Cover, R. L. Matsueda, and C. E. Kubrin. 2006. "Neighbourhood Context and Neighbouring Ties." *City and Community* 5: 363–85.

Harding, D. 2003. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on High School Dropout and Teenage Pregnancy." *American Journal of Sociology* 109 (3): 676–719.

Harris, D. 1999. "'Property Values Drop When Blacks Move In Because …': Racial and Socioeconomic Determinants of Neighborhood Desirability." *American Sociological Review* 64: 461–79.

Hedman, L. 2011. "The Impact of Residential Mobility on Measurements of Neighbourhood Effects." *Housing Studies* 26 (4): 501–19.

Hedman, L., and M. van Ham. 2012. "Understanding Neighbourhood Effects: Selection Bias and Residential Mobility." In *Neighbourhood Effects Research: New Perspectives,* edited by M. van Ham, D. Manley, N. Bailey, L. Simpson, and D. Maclennan (79–99). Dordrecht, Netherlands: Springer.

Hedman, L., M. van Ham, and D. Manley. 2011. "Neighbourhood Choice and Neighbourhood Reproduction." *Environment and Planning A* 43 (6): 1381–99.

Hedman, L. 2013. "Moving Near Family? The Influence of Extended Family on Neighbourhood Choice in an Intra-Urban Context." *Population, Space and Place* 19: 32–45.

Hedman, L., and G. Galster. 2013. "Neighbourhood Income Sorting and the Effects of Neighbourhood Income Mix on Income: A Holistic Empirical Exploration." *Urban Studies* 50: 107–127.

Ioannides, Y. M., and J. E. Zabel. 2008. "Interactions, Neighbourhood Selection and Housing Demand." *Journal of Urban Economics* 63: 229–52.

Jacob, B. 2004. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *American Economic Review* 94: 233–58.

Jencks, C., and S. Mayer. 1990. "The Social Consequences of Growing Up in a Poor Neighborhood." In *Inner-City Poverty in the United States,* edited by L. Lynn and M. McGeary. Washington, DC: National Academy Press.

Kauppinen, T. 2007. "Neighbourhood Effects in a European City: Secondary Education of Young People in Helsinki." *Social Science Research* 36: 421–44.

Kearns, A., and A. Parkes. 2003. "Living In and Leaving Poor Neighbourhoods in England." *Housing Studies* 18 (6): 827–51.

Kling, J., J. Liebman, and L. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Kling, J. R., J. Ludwig, and L. F. Katz. 2005. "Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment." *Quarterly Journal of Economics* 120 (1): 87–130.

Lee, B. A., R. S. Oropresa, and J. W. Kanan. 1994. "Neighborhood Context and Residential Mobility." *Demography* 31 (2): 249–70.

Lee, B., and P. Wood. 1991. "Is Neighborhood Racial Succession Place-Specific?" *Demography* 28 (1): 21–40.

Leventhal, T., and J. Brooks-Gunn. 2000. "The Neighbourhoods They Live In: The Effects of Neighborhood Residence on Child and Adolescent Outcomes." *Psychological Bulletin* 126 (2): 309–37.

Lim, U., and G. Galster. 2009. "The Dynamics of Neighborhood Property Crime Rates." *Annals of Regional Science* 43 (4): 925–45.

Lu, M. 1998. "Analyzing Migration Decisionmaking: Relationships between Residential Satisfaction, Mobility Intentions, and Moving Behaviour." *Environment and Planning A* 30: 1473–95.

Ludwig, J., H. Ladd, and G. Duncan. 2001. "The Effects of Urban Poverty on Educational Outcomes: Evidence from a Randomized Experiment." In *Brookings-Wharton Papers on Urban Affairs,* edited by W. Gale and J. R. Pack (147–201). Washington, DC: Brookings Institution.

Ludwig, J., J. B. Liebman, J. R. Kling, G. J. Duncan, L. F. Katz, R. C. Kessler, and L. Sanbonmatsu. 2008. "What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment?" *American Journal of Sociology* 114 (1): 144–88.

Manski, C. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14 (3): 115–36.

Molina, I. 1997. *Stadens rasifiering. Etnisk boendesegregation i folkhemmet.* Geografiska regionstudier 32. Uppsala, Sweden: Uppsala University.

Murray, M. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20: 111–32.

Musterd, S., and R. Andersson. 2006. "Employment, Social Mobility and Neighbourhood Effects: The Case of Sweden." *International Journal of Urban and Regional Research* 30 (1): 120–40.

Musterd, S., G. Galster, and R. Andersson. 2012. "Temporal Dimensions and Measurement of Neighbourhood Effects." *Environment and Planning A* 44: 605–27.

Musterd, S., W. Ostendorf, and S. de Vos. 2003. "Neighbourhood Effect and Social Mobility: A Longitudinal Analysis." *Housing Studies* 18 (6): 877–92.

Oreopoulos, P. 2003. "The Long-Run Consequences of Living in a Poor Neighborhood." *Quarterly Journal of Economics* 118: 1533–75.

Orr, L., J. Feins, R. Jacob, E. Beecroft, L. Sanbonmatsu, L. Katz, J. Liebman, and J. Kling. 2003. *Moving to Opportunity: Interim Impacts Evaluation.* Washington, DC: US Department of Housing and Urban Development.

Pais, J. F., S. South, and K. Crowder. 2009. "White Flight Revisited: A Multiethnic Perspective on Neighborhood Out–Migration." *Population Research and Policy Review* 28 (3): 321–46.

Parkes, A., A. Kearns, and R. Atkinson. 2002. "What Makes People Dissatisfied with Their Neighbourhoods?" *Urban Studies* 39 (13): 2413–38.

Permentier, M., M. van Ham, and G. Bolt. 2007. "Behavioural Responses to Neighbourhood Reputation." *Journal of Housing and the Built Environment* 22: 199–213.

Permentier, M., M. van Ham, and G. Bolt. 2009. "Neighbourhood Reputation and the Intention to Leave the Neighbourhood." *Environment and Planning A* 41 (9): 2162–80.

Propper, C., K. Jones, A. Bolster, S. Burgess, R. Johnston, and R. Sarker. 2005. "Local Neighbourhood and Mental Health: Evidence from the UK." *Social Science and Medicine* 61 (10): 2065–83.

Quercia, R., and G. Galster. 2000. "Threshold Effects and Neighborhood Change." *Journal of Planning Education and Research* 20 (2): 146–62.

Quillian, L. 2002. "Why Is Black–White Residential Segregation So Persistent? Evidence on Three Theories from Migration Data." *Social Science Research* 31: 197–229.

Robinson, D. 2002. "Missing the Target? Discrimination and Exclusion in the Allocation of Social Housing." In *'Race', Housing and Social Exclusion,* edited by P. Somerville and A. Steele. London: Jessica Kingsley Publishing.

Roseman, C.C. 1983. "A Framework for the Study of Migration Destination Selection." *Population and Environment* 6 (3): 151–65.

Rosenbaum, J. 1995. "Changing the Geography of Opportunity by Expanding Residential Choice: Lessons from the Gautreaux Program." *Housing Policy Debate* 6 (1): 231–69.

Ross, C.E., and J. Mirowsky. 2008. "Neighborhood Socioeconomic Status and Health: Context or Composition?" *City and Community* 7 (2): 163–79.

Ross, S.J., and M.A. Turner. 2005. "Housing Discrimination in Metropolitan America: Explaining Changes between 1989 and 2000." *Social Problems* 52: 152–80.

Rothenburg, J., G. Galster, R. Butler, and J. Pitkin. 1991. *The Maze of Urban Housing Markets: Theory, Evidence and Policy.* Chicago: University of Chicago Press.

Sampson, R.J., and W.B. Groves. 1989. "Community Structure and Crime: Testing Social-Disorganization Theory." *American Journal of Sociology* 94 (4): 774–802.

Sampson, R., J. Morenoff, and T. Gannon-Rowley. 2002. "Assessing 'Neighborhood Effects': Social Processes and New Directions in Research." *Annual Review of Sociology* 28: 443–78.

Sampson, R., and P. Sharkey. 2008. "Neighborhood Selection and the Social Reproduction of Concentrated Racial Inequality." *Demography* 45 (1): 1–29.

Sampson, R., P. Sharkey, and S. Raudenbush. 2008. "Durable Effects of Concentrated Disadvantage on Verbal Ability among African-American Children." *Proceedings of the National Academy of Sciences* 105: 845–52.

Sari, F. 2012. "Analysis of Neighbourhood Effects and Work Behaviour: Evidence from Paris." *Housing Studies* 27 (1): 45–76.

Schelling, T. 1971. "Dynamic Models of Segregation." *Journal of Mathematical Sociology* 1: 143–86.

Schwartz, H. 2010. *Housing Policy Is School Policy: Economically Integrative Housing Promotes Academic Achievement in Montgomery County, MD.* New York: Century Foundation.

Skifter Andersen, H. 2003. *Urban Sores: On the Interaction between Segregation, Urban Decay and Deprived Neighbourhoods.* Aldershot, United Kingdom: Ashgate.

Smith, S. J. 1989. *The Politics of 'Race' and Residence: Citizenship, Segregation and White Supremacy in Britain.* Cambridge, United Kingdom: Polity Press.

South, S. J., and K. D. Crowder. 1997. "Escaping Distressed Neighbourhoods: Individual, Community and Metropolitan Influences." *American Journal of Sociology* 102 (4): 1040–84.

South, S. J., and K. D. Crowder. 1998. "Leaving the 'Hood': Residential Mobility between Black, White and Integrated Neighbourhoods." *American Sociological Review* 63: 17–26.

Speare, A. 1974. "Residential Satisfaction As an Intervening Variable in Residential Mobility." *Demography* 11: 173–88.

Sykes, B., and S. Musterd. 2011. "Examining Neighbourhood and School Effects Simultaneously: What Does the Dutch Evidence Show?" *Urban Studies* 48 (7): 1307–31.

Taub, R. P., D. G. Taylor, and J. D. Dunham. 1984. *Paths of Neighbourhood Change: Race and Crime in Urban America.* Chicago: University of Chicago Press.

Tienda, M. 1991. "Poor People and Poor Places: Deciphering Neighborhood Effects on Poverty Outcomes." In *Macro-Micro Linkages in Sociology,* edited by J. Huber. Newbury Park, CA: Sage Publications.

Turley, R. 2003. "When Do Neighborhoods Matter? The Role of Race and Neighborhood Peers." *Social Science Research* 32: 61–79.

van Ham, M., and W. A. V. Clark. 2009. "Neighbourhood Mobility in Context: Household Moves and Changing Neighbourhoods in the Netherlands." *Environment and Planning A* 41: 1442–59.

van Ham, M., and P. Feijten. 2008. "Who Wants to Leave the Neighbourhood? The Effect of Being Different from the Neighbourhood Population on Wishes to Move." *Environment and Planning A* 40: 1151–70.

van Ham, M., and D. Manley. 2010. "The Effect of Neighbourhood Housing Tenure Mix on Labor Market Outcomes: A Longitudinal Investigation of Neighbourhood Effects." *Journal of Economic Geography* 10 (2): 257–82.

Vartanian, T. P., and P. M. Gleason. 1999. "Do Neighbourhood Conditions Affect High School Dropout and College Graduation Rates?" *Journal of Socio-Economics* 28: 21–41.

Vartanian, T. P., P. Walker Buck, and P. Gleason. 2007. "Intergenerational Neighborhood-Type Mobility: Examining Differences between Blacks and Whites." *Housing Studies* 22 (5): 833–56.

Wheaton, B., and P. Clarke. 2003. "Space Meets Time: Integrating Temporal and Contextual Influences on Mental Health in Early Adulthood." *American Sociological Review* 68 (5): 680–706.

Wilson, W. J. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press.

Wodtke, G., D. Harding, and F. Elwert. 2011. "Neighborhood Effects in Temporal Perspective: The Impact of Long-Term Exposure to Concentrated Disadvantage on High School Graduation." *American Sociological Review* 76 (5): 713–36.

Wolf, E. 1963. "The Tipping-Point in Racially Changing Neighborhoods." *Journal of the American Institute of Planners* 29 (3): 217–22.

ESSAY

# Beyond Mapping

## *Spatial Analytics and Evaluation of Place-Based Programs*

Julia Koschinsky

Over the past decade, community change efforts have increasingly used geographically based community information (Brown 2010), primarily to visualize data in a mapping format. The collection of such geo-referenced data has been advanced by several community data initiatives, including, for instance, the efforts organized through the Urban Institute's National Neighborhood Indicators Partnership or projects in Chicago such as MDRC's neighborhood trajectory database for the New Communities program evaluation, RW Ventures's Dynamic Taxonomy Project database, or Sampson's (2012) analysis of neighborhood effects using Project on Human Development in Chicago Neighborhoods data.

This essay builds on these developments and introduces explicitly spatial concepts, methods, and tools to the evaluation of place-based programs. Some of these spatial approaches have been applied in particular place-based evaluations before, but the use of a spatial analytical perspective beyond mapping geographic data is less common. This perspective has been neglected in the field of evaluation in general. In contrast to other research areas with exponential growth in the application of spatial methods and tools in the last decade, spatial analysis beyond mapping is less well-known or used in evaluation practice and research (Koschinsky 2013). Existing evaluation research and practice tend not to go beyond the standard functionalities of mapping tools [such as geographic information systems (GISs); see, for example, Renger et al. (2002)]. There is an opportunity to benefit from recent research that contextualizes GIS tools in the larger and more theoretically grounded field of geographic information science (National Research Council 2006). Hence this essay seeks to illustrate how such a spatial perspective, which has been applied fruitfully in other disciplines, can also reveal important spatial dimensions of place-based initiatives and evaluations that are likely to be missed with a more narrow focus on spatial tools.

A more systematic spatial framing of place-based evaluations reveals relevant gaps in the current use of geographic information. One such gap is a predominant focus on visualizing how geographic areas compare in regard to an indicator of interest (so-called spatial heterogeneity). An opportunity exists to extend this focus on heterogeneity to interactions between areas and individuals, such as through multiplier effects, displacement effects, or other forms of so-called spatial dependence. Spatial interaction or feedback effects can not only be detected through various cluster and hotspot methods, but also modeled more explicitly through multivariate spatial econometric methods (briefly discussed below). These methods go beyond the general practice of eyeballing concentrations of values on a map by quantifying spatial correlations and testing against a null hypothesis of spatial randomness. Finally, a trend within geographic information science away from desktop applications with manual data inputs and static map outputs toward more flexible cyber-GIS frameworks presents an opportunity in evaluation research to start envisioning the integration of spatial data analysis within learning systems that are part of decision workflows, with near-real-time data inputs and more automated visualization and analysis of spatial data.

The essay begins with a brief discussion of spatial analysis and place-based evaluation, followed by an overview of some key spatial concepts of a spatial perspective [modifiable areal unit problem (MAUP), spatial dependence, and spatial heterogeneity] and how they apply to the evaluation of place-based programs. This discussion is followed by an overview of spatial data and selected methods and tools of particular relevance to place-based evaluation. The essay concludes with an identification of three research gaps at the intersection of spatial analysis and evaluation.

## Spatial Analysis and Place-Based Evaluations

Spatial thinking, methods, and tools make sense when the null hypothesis that areawide averages are generally representative of local characteristics does not hold. Rather, spatial thinking "uses space to integrate and structure ideas" (National Research Council 2006, x). Spatial tools implement methods for spatial data analysis, which is "the formal quantitative study of phenomena that manifest themselves in space. This implies a focus on location, area, distance and interaction" (Anselin 1990, 64). Spatial concepts and methods are relevant in instances when the *end of geography* is a less accurate description than Tobler's (1970) first law of geography that "everything is related to everything else, but near things are more related than distant things." And they can add insights when individuals are not assumed to act as isolated economic agents but are influenced by social interaction effects (Durlauf 2004) and neighborhood effects (Sampson 2012). What matters here is that traditional statistical and economic assumptions of independence are replaced by an explicit focus on and treatment of spatially constrained interdependence.

Spatial thinking, methods, and tools are congruent with place-based efforts because, at a minimum, what these programs have in common is a focus on a shared delineated location. More broadly, "placed-based approaches are collaborative means to address complex social-economic issues through interventions defined at a specific geographic scale" (Cantin, cited in Bellefontaine and Wisener 2011, 7). Examples of place-based approaches in the United States that particularly emphasize comprehensive approaches and community building include comprehensive community initiatives and other community change efforts (Kubisch et al. 2010).

Space becomes a platform for integration across academic disciplines, such as computer science, economics, statistics, and geography, as illustrated by the subfields of geo-computation, geo-statistics, and spatial econometrics (Goodchild and Janelle 2004). It also becomes a platform for coordinating and evaluating interventions across sectors and governmental levels, domains, and departments such as employment, housing, education, and crime. Space invites data integration across jurisdictions and scales through GISs that layer data of multiple types and levels of aggregation. The recent focus among funding agencies on collective impact illustrates how funders who are investing in the same geographic areas can commit to collaborate and jointly monitor programs in these areas (e.g., in the 300-organization Strive Partnership to improve children's educational and employment outcomes in Cincinnati, Ohio). This approach contrasts with silo practices in which each funder designs and evaluates its own programs separately, even when programs pertain to the same area.

The assumption that different areas warrant different program designs or expectations of how impacts work contrasts with one-size-fits-all, blanket approaches that generally do not take spatial considerations into account. As such, place-based approaches recognize so-called spatial effects. That is, problems related to poverty, youth unemployment, housing shortages, or school quality vary by region (spatial heterogeneity), and even within regions, they cluster in particular neighborhoods (spatial dependence).

## Some Key Spatial Concepts for Place-Based Evaluation

### Modifiable Areal Units and Target Area Boundaries

Data analysis for place-based evaluations is usually based on data for observations summarized at the neighborhood and target area levels, often because Census data and other data are only available in aggregate format or because it can be easier to visualize patterns for known areas. This practice makes the evaluation of place-based initiatives vulnerable to MAUP (Openshaw 1984). The boundaries of areal units are modifiable for two reasons: multiple scales (it is often unclear how many zones to use) and aggregation (it is equally unclear how to group the zones). For instance, housing parcels could be grouped into zones at a large number of scales such as (from presumably smallest to largest) census blocks, block groups, tracts, Zip Codes, perceived neighborhood boundaries, housing submarkets, and so on. How to group these zones into target areas for place-based initiatives is also associated with ambiguity.

This uncertainty is problematic because descriptive and statistical results (including spatial statistical results) are likely to vary significantly depending on which scale and aggregation zones are used (gerrymandering is a classic example of how election outcomes are influenced by how households are aggregated). In other words, outcomes are likely to change depending on which and how many zones households are assigned to. In a classic quantitative analysis, Openshaw and Taylor (1979) concluded that the size of a correlation coefficient expresses a relationship between variables of interest that changes with scale and aggregation levels, resulting in "a million or so correlation coefficients." In other words, correlation coefficients are modifiable with the areal unit. Specifically, their size tends to be inversely related to the number of zones; that is, they often increase as the number of geographical areas decreases (Openshaw 1984).

MAUP has several implications for the analysis of data in evaluations of place-based initiatives. For one, it highlights the potential dependence of mapping and statistical results on the areal units chosen. Even if target or neighborhood boundaries are bound by community preferences and/or the availability of other sources (such as census boundaries, Zip Codes, school districts, or police beats), it is important to be aware of MAUP when there are choices about how existing units are grouped or aggregated. This awareness may become more relevant as large margins of error associated with key neighborhood-level poverty-related estimates of the American Community Survey make indicators from alternative address-based sources with flexible aggregation options (such as point-based sales prices) more attractive. As more address-level data become available, alternative spatial units of analysis also become available as options to consider. For instance, in crime- and transit-related analyses, points have recently been aggregated to street segments to obtain more spatially focused results (Weisburd, Groff, and Yang 2012). In one such analysis, Weisburd, Morris, and Groff (2009) found that 50 percent of all juvenile crimes in Seattle were committed on less than 1 percent of street segments. Such a refinement helps target intervention efforts better than the standard kernel density hotspot maps with vague boundaries.

MAUP also raises the questions of if and how a scale mismatch (i.e., a mismatch in boundaries from different data sources) might influence results, and how this problem can be addressed methodologically (see the section on spatial modeling). Substantively, these questions are related to the fact that the dynamics that influence local poverty concentrations operate at different scales. For example, housing and labor market dynamics are regional or larger, but school district outcomes might be more local. Weissbourd, Bodini, and He (2009, 2) estimate that over one-third of neighborhood change is related to trends at a larger regional scale. Two related questions are what other spatially targeted efforts are underway for a particular target area and how these efforts might affect the processes and outcomes of the evaluated target area. Examples include health service areas, neighborhood planning zones, business improvement districts, transportation plans, and consolidated planning areas.

MAUP is also relevant in the analysis of neighborhood effects, which can be thought of as an analysis in which one scale (community outcomes) influences another (individual outcomes). How to delineate and measure neighborhoods in this research can have important impacts on results. A good example in this context is the federal Moving to Opportunity experiment, which was designed to test, among other things, if public housing residents in concentrated poverty neighborhoods would improve employment outcomes as a result of moving to lower-poverty neighborhoods. Although the Moving to Opportunity interim evaluation (Goering and Feins 2003) did not find any significant impacts in this regard, Clampet-Lundquist and Massey (2008) reanalyzed the data and found significant positive impacts after controlling for the time residents actually spent in lower-poverty neighborhoods after moving frequently.

The problem of modifiable zone boundaries is also related to the question of how constant these boundaries remain over time and how much the composition of residents changes. For instance, if programs in target zones are designed to improve individual outcomes (e.g., through job training, education, or health services), what is the extent of in- and out-migration of the target area, and how does this migration affect program theory, design, and outcomes? Although place-based work is often based in areas of concentrated poverty to reach residents affected by it, what proportion of residents who are employed by businesses or using services in the target area actually lives within (or outside of) the target area? Relevant research includes the assessment by Coulton, Theodos, and Turner (2009) of the impact of mobility on neighborhood change in the context of evaluating community initiatives (Kubisch et al. 2010, chapter 7).

Research on neighborhood change has traditionally focused on fixed neighborhood boundaries with changing socioeconomic indicators. In an example of recent spatial analysis that seeks to address MAUP and resident mobility, Rey et al. (2011) identify neighborhood change not only based on such socioeconomic indicators but also based on neighborhood boundaries. The authors develop a new neighborhood change index that is based on a spatially constrained cluster algorithm that generates new neighborhood areas based on value and area similarity. In contrast to the assumption that neighborhood boundaries remain relatively constant over time, they find that change in the spatial extent of neighborhoods "appears to be the rule rather than the exception because two thirds of the census tracts [were] found to experience some reconfiguration in what constitute[d] their neighborhood set between 1990 and 2000" (Rey et al. 2011, 61). Denser inner-city areas (which are more likely to host place-based programs) experienced more of these neighborhood changes than outlying areas during this period. The notion that neighborhoods are not fixed entities is congruent with findings of RW Ventures's Dynamic Neighborhood Taxonomy project that 70 percent of households in the four cities they evaluated moved at least once during a 10-year period (Weissbourd et al. 2009, 2).

## Beyond Mapping Spatial Heterogeneity: Introducing Spatial Dependence

Spatial analysis is designed to address spatial effects, which have two dimensions: spatial dependence and spatial heterogeneity (Anselin 1988). Like Tobler's first law of geography, spatial autocorrelation (spatial correlation of the same variable between different areas) operationalizes the concept of spatial dependence. Spatial autocorrelation exists when there is a functional relationship between the values in one location and those in another location. In other words, data aggregated at an area level are autocorrelated when what happens in one location depends on what happens in nearby locations. The first law's notion that nearby things are more related implies the existence of spatial threshold effects [see also Quercia and Galster (2000)]. Spatial autocorrelation can be measured on a global level (as the extent of overall clustering in the study area) or at a local level (as local clusters and spatial outliers within the study area). Positive spatial autocorrelation refers to the correlation of similar values of the same variable in space (e.g., high poverty rates surrounded by other high poverty rates or low home values surrounded by other low home values). Negative spatial autocorrelation identifies the correlation of dissimilar values in space, such as low values surrounded by high values or vice versa (e.g., in gentrifying areas).

In contrast to spatially autocorrelated data, in which values are dependent on each other by relative location, spatially heterogeneous data differ by location, but these differences do not result from the interaction between locations. For instance, if areas A and B are different but independent, they are spatially heterogeneous. If area A is different because of area B (and vice versa, as B is A's neighbor), they are spatially dependent. Spatial heterogeneity refers to the absence of spatial stationarity, which is related to structural changes in the data that are related to location. Spatial stationarity assumes homogeneity across space, for instance, as captured by one global parameter that equally characterizes all subregions within a study area (as assumed in nonspatial ordinary least squares (OLS) regression).

Empirically, it is difficult to distinguish spatial dependence and spatial heterogeneity, especially with cross-sectional data. For instance, two census tracts' poverty rates can be correlated because of peer effects between residents in both tracts, shared economic conditions in both tracts (such as a factory closing that affected both or cheaper land values in both), and interactions between these dynamics. In spatial point pattern analysis, the difference between "true and apparent contagion" (with origins in epidemiology) is analogous to that between spatial dependence and spatial heterogeneity in the analysis of spatial areas. In a process of true contagion, two people end up with a similar characteristic as a result of an interaction in which the characteristic was transmitted or shared. In a process of apparent contagion, two people have similar characteristics, but this similarity is independent of any interaction they might have. True and apparent contagions are observationally equivalent (i.e., the same observed results could be reached through different unobserved processes). Since the underlying spatial process that drives the observed spatial pattern is unobserved, true and apparent contagion cannot be distinguished in a cross-sectional context without additional information, such as additional spatially independent observations or additional observations from different time periods.

Spatial feedback effects that exist under conditions of true contagion violate the lack of interaction assumption whereby treatments for one person do not influence others' outcomes. This logic of independence was based on randomized controlled trials designed to isolate impacts of single interventions. However, part of what makes spatially clustered problems "wickedly complex" is that they have interacting causes (Bellefontaine and Wisener 2011). Here individual-level outcomes are often dependent, as in spillover effects on a single outcome or between different outcomes (Bloom 2005). In spillovers among single outcomes, the outcome of one or more people influences that of others (e.g., in peer effects, when employment increases as a result of networking with employed neighbors). Spillovers between different outcomes occur when one outcome influences a different one, as in neighborhood effects when the chance that a child does well in school would be lower in a neighborhood of concentrated poverty than in a more affluent one.

At the aggregate level, the concept of spatial dependence can focus attention on potential interactions between a target area and neighboring areas, including intended and unintended consequences. For instance, one such question is whether a place-based approach has positive spatial externalities (such as economic multiplier effects in neighboring areas) or negative spatial externalities (which may occur if a crime reduction effort in one precinct ends up displacing crime to nearby precincts,[1] or if subsidies to target area firms end up reducing growth in similar nearby firms). Another question focuses on the extent to which outcomes within a target area depend on interactions with other areas given that spatial target areas are open systems (i.e., they constantly interact with other areas). For instance, do target areas with stronger economic linkages to other areas perform better economically? Such linkages can pertain to business networks or spatial infrastructure such as roads, public transit, railways, or broadband that connect areas in physical or virtual space. RW Ventures's Dynamic Neighborhood Taxonomy is an example of a tool to help contextualize neighborhoods and cities that are part of place-based efforts within the broader regional economy. At a household level, a question in this context is to what extent residents' social networks are defined by a target area and what the consequences are for place-based program goals (Livehoods.org is an interesting recent application to delineate neighborhood boundaries in near real time based on social media use in neighborhoods).

The predominant focus in place-based approaches and evaluations that have used spatial data (as in program evaluation generally) has been on spatial heterogeneity, not on spatial dependence. Maps in this context often visualize the extent of heterogeneity of key indicators within a target area and how a target area compares with the region. Interpretation of these maps often involves statements such as "as one can see, poverty is concentrated in these areas." In cases of strong concentration, such as highly segregated cities, such statements will hold against a null hypothesis of spatial randomness. However, in many common but less extreme cases, people tend to overdetect clusters where none exist. In these cases, analytical (as opposed to visual) spatial methods are helpful in distinguishing spurious clusters from statistically significant ones. Hence this essay argues for a broader application of spatial concepts and methods, beyond spatial heterogeneity, to include spatial significance testing and modeling.

## Spatial Data, Methods, and Tools for Evaluation

One of the key developments in evaluating place-based initiatives that address wickedly complex problems has been a move away from measuring the impacts of these initiatives within a causal framework designed for single causal effects. Instead of asking the extent to which programs caused an outcome, evaluation approaches emerged that ask whether how programs actually work matches the expectations of how and why they are supposed to work [popularized in the theory of change (Weiss 1995), which itself is seen as emergent (Patton 2011)]. Here causal attribution of program impacts is replaced by "contribution analysis" (Mayne 2001), in which the question is whether program investments are correlated with community change.[2] Recent research on neighborhood effects also addresses the problem of multiple connected causes through a theory of "contextual causality" (Sampson 2012).[3] This shift toward contribution analysis has implications for spatial data analysis because it implies a greater emphasis on, for instance, tracking indicators of contextual factors and identifying statistical correlations between program outputs and community outcomes rather than seeking to isolate causal connections between outputs and outcomes.

As methods vary by data type, an overview of spatial data types is followed by a brief discussion of selected spatial methods and tools that are relevant to place-based evaluations.

## Spatial Data

Classic taxonomies of spatial data distinguish between area data, event points, spatially continuous data, and spatial interaction or flow data (Bailey and Gatrell 1995).[4] Area data consist of discrete areas with attribute variation between areas or discrete points (see the MAUP section above). A common way to define which areas are neighbors is a so-called spatial weights matrix (often identified as W in equations), which identifies neighboring areas for a given area based on connectivity criteria, such as queen border contiguity, nearest neighbors, or bands of network distance. Event points are points whose location is subject to uncertainty (such as crimes). Bandwidths around points are often used to define neighboring points. Spatially continuous or geo-statistical data describe phenomena that are continuously distributed over space and are measured at sample points (such as contaminated soil). Spatial interaction or flow data characterize movement between an origin and a destination, such as in social networks or walkable access to school. See Radil, Flint, and Tita (2010) for an application and LeSage and Pace (2010) for spatial econometrics methods for flow data.

The focus in this essay is on area data, because much socioeconomic data are collected for fixed administrative units (such as census tracts). In this context, spatial econometric methods have been developed to incorporate spatial effects (Anselin 1988). Event point data are frequently applied in crime analysis and epidemiology (Waller and Gotway 2004). The use of continuous spatial data primarily characterizes applications in the natural sciences, such as geology, biology, or forestry, and is based on geophysical data and geo-statistical methods (Bailey and Gatrell 1995; Cressie 1991; Cressie and Wikle 2011).

Although the focus in this essay is on geographic space, spatial connectivity can be conceptualized nongeographically (e.g., in terms of social connectivity). Examples include Hanssen and Durland's[5] application of social network analysis to relate denser math teacher communication networks to improved student test scores or the assessment by Radil et al. (2010) of gang violence in geographic and social network space. Perceived distances and neighborhood boundaries are another example of non-Euclidean space (Coulton et al. 2001).

## Spatial Methods and Tools

This section highlights some standard and new spatial analysis methods that I believe are particularly relevant for the evaluation of place-based initiatives in the context of this essay. Among the myriad of spatial methods that exist, I selected a few that can generate interesting insights without being too complicated to explain, at least conceptually, to stakeholders without statistical backgrounds.[6] Selected tools to implement these methods are also mentioned, with an emphasis on programs that are free and/or open source, user friendly, under active development, and with larger user bases. However, these selections are necessarily subjective and biased toward some of the free and open-source software development efforts I am affiliated with.[7] For a broader spatial analysis software overview, see reviews by Fischer and Getis (2010) and Anselin (2012).[8]

## Cluster Detection Methods

The most common techniques for identifying spatial concentrations of values beyond eyeballing map patterns are methods that detect statistically significant clusters (called hotspots in some fields). These methods can operationalize the concepts of spatial dependence (spatial autocorrelation) discussed above. One of the popular groups of contiguity-based cluster methods are so-called local indicators of spatial association (LISAs) (Anselin 1995), which identify whether an area has statistically similar values compared with its neighbors and, when summarized across all areas, is proportionate to a global indicator of spatial association. Popular LISA statistics are local Moran's I and Gi/Gi*, which can be used to map hotspots, coldspots, and/or spatial outliers (areas with values that are inversely related to those of their neighbors, as explained above) (Anselin, Sridharan, and Gholston 2007). Programs in which these maps can be created include OpenGeoDa, R (R-Geo packages, such as GeoXP), the Python-based Spatial Analysis Library PySAL (all free and open source), and Esri's spatial statistics toolbox.

As mentioned above, street segments have frequently been used as units of analysis to which points are aggregated. Several recent methods extend existing cluster techniques to networks, including kernel density mapping (Okabe, Satoh, and Sugihara 2009); local Ripley's K function (Okabe and Yamada 2001); and LISAs (local Moran's I and Gi/I*) (Yamada and Thill 2010). These improved methods allow analysts to generate cluster maps at the street segment level that can be visualized through varying street colors, widths, or heights. Existing software to apply some or all of these methods include SANET (for ArcGIS), GeoDaNet, and PySAL 1.4. These programs also allow users to compute street network distances between origins and destinations;[9] an analyst could determine, for example, if services are within reach of target area residents. GeoDaNet also generates spatial weights based on network distances. Another program, the Urban Network Analysis Toolbox for ArcGIS (recently released by the Massachusetts Institute of Technology's City Form Lab), enables the analysis of point patterns using graph methods with street networks.

Several recent exploratory methods add a time dimension to the traditional cross-sectional spatial cluster methods to detect local clusters across space and time. A popular local test for detecting space–time hotspots of event points is the space–time permutation scan by Kulldorff et al. (2005) implemented in the SaTScan software (Kulldorff 2010). This method detects events (such as crimes) that occurred in a similar location at a comparable time and tests the significance of these space–time hotspots by using a Monte Carlo permutation approach [for a comprehensive overview of statistical analysis of spatial points patterns, see Diggle (2003)].

Rey (2001) integrated LISAs with classic Markov transition matrices to test if joint transitions of an area and its neighbors remain significant for a given set of time periods. One of the questions this method can address is if area-based hotspots (e.g., of poverty, crime, or house prices) persist over time. The LISA Markov is implemented in PySAL.[10]

A different approach to finding clusters is based on so-called spatially constrained clustering algorithms (Duque, Anselin, and Rey 2012). This approach uses optimization methods to group spatial units with similar characteristics into new regions. An application of this method to measure neighborhood change based on changing boundaries and social composition was highlighted in the MAUP discussion. The clusterPy library (GeoGrouper with a graphical user interface) and PySAL's regionalization code implement these methods.

Cluster maps can be linked to other spatial and nonspatial representations of data (such as parallel coordinate plots, scatterplots, conditional plots, histograms, or trend graphs) in open-source desktop programs such as OpenGeoDa, web platforms such as Weave (WEb-based Analysis and Visualization Environment), or through customizing visualization libraries.[11]

## Spatial Modeling

The cluster techniques discussed in the previous section are exploratory in the sense that they generally only allow for spatial pattern detection in a uni- or bivariate context or when multiple variables are considered together; this pattern detection is done visually rather than statistically. As such, these techniques are best suited for generating rather than testing hypotheses. To control for a larger number of variables in a model that tests hypotheses related to an underlying theory, multivariate spatial models are needed. If the goal is to model spatial heterogeneity (i.e., spatial variation among regions whose similarity is not due to interaction), specialized spatial econometric methods are not needed. Discrete differences between subregions are often modeled through so-called spatial regimes with separate coefficients that can be tested for equivalence with spatial Chow tests (Anselin 1988). Further, geographically weighted regression has been popularized by Fotheringham, Brunsdon, and Charlton (2002) to estimate continuous forms of spatial variation through locally varying coefficients.

However, in the context of spatial dependence, an intersection between the fields of econometrics, program evaluation, and spatial analysis continues to be missing. This lack persists despite the fact that recent research has started to strengthen the bridge between econometrics and program evaluation and the older existing connection between the fields of statistics and evaluation.[12] One of the few research projects that has started to connect spatial statistical methods and program evaluation leverages spatial Bayesian approaches (Banerjee, Carlin, and Gelfand 2004; Cressie and Wikle 2011). For instance, Verbitsky (2007)[13] uses spatial Bayesian hierarchical modeling to evaluate whether Chicago's community policing programs displaced crime to neighboring areas.[14] Verbitsky Savitz and Raudenbush (2009) test the performance of a spatial empirical Bayesian estimator in an application related to neighborhood collective efficacy, which is central to research on neighborhood effects (Sampson 2012). The spBayes package in R and WinBUGS can be used to estimate spatial Bayesian models.

Spatial autocorrelation as discussed above (e.g., in the form of spillover effects) can be associated with the dependent variable (Wy), independent variables (Wx), and the error term (We) of a model. Spatial estimators are not required when the values of an area and its neighbors are correlated for the independent variables. However, the presence of spatial autocorrelation in the dependent variable or error term violates standard assumptions in classic statistics and can thus result in inconsistent and/or biased estimates. Hence the correction of these problems requires specialized techniques beyond traditional nonspatial methods. Two standard linear spatial regression models that account for such spatial dependence are called *spatial lag* and *spatial error* models in the spatial econometrics literature (Anselin 1988).

To determine if spatial effects are present, spatial diagnostic tests can be applied (e.g., spatial Lagrange multiplier tests to the residuals of a standard nonspatial OLS model). These tests suggest if spatial lag and spatial error models are a better fit for a given dataset. In spatial lag model specifications, the dependent variable is not only a function of the independent variables and the error term, but also of the average value of the dependent variable of the neighboring observations of any given observation. This model can be consistent with spatial multiplier effects. In spatial error models, the error term is spatially correlated, which can, for example, be related to a mismatch in scales (e.g., between the spatial extent of concentrated poverty and the administrative boundary for which data are available). Recent advances in the estimators of spatial lag and error models also control for nonconstant error variance (so-called heteroskedasticity) (Arraiz et al. 2010) or adjust OLS estimates for spatial correlation and heteroskedasticity (Kelejian and Prucha 2007). Versions of these methods have recently been implemented in Stata (spivreg package), R (sphet package), PySAL 1.3, and GeoDaSpace. The Arc_Mat toolbox in MatLab (Liu and LeSage 2010) also contains spatial econometric estimators in addition to spatial exploratory tools.

A key potential contribution of these cross-sectional linear models is that they allow for the measurement of the correlation between contextual factors and outcome variables, an ability that can inform the question of whether a program contributed to an observed change in a community. However, an important research gap exists in space–time model specifications, estimators, and tools that more comprehensively account for differences in outcomes before and after programs start to operate. This gap includes spatial estimators for difference-in-difference–type models. Existing approaches add so-called spatial fixed effects, that is, spatially lagged independent variables (Wx) are added to a standard difference-in-difference specification (Galster et al. 2004; Ellen et al. 2001). However, there are some methodological problems with this approach (Anselin and Arribas-Bel 2011). A few examples of other recent spatial approaches relevant to evaluations include spatial seemingly unrelated regression models that account for spatial dependence by incorporating either a spatial lag or spatial error term (Anselin 1988). When time periods are pooled so a model of outcomes and related factors can be estimated before and after an initiative started, then a test (Wald) can be applied to assess significant differences in estimates for the two time periods. However, selection effects are often not accounted for here. In a randomized controlled trial context, Bloom (2005) advanced solutions for measuring program impacts in group randomization designs in which spillover effects between the same or different outcomes are present. An example of integrating spatial concepts with a classic quasi-experimental design is Fagan and MacDonald's (2011)[15] incorporation of information about locations, neighboring areas, and time in their regression discontinuity design. Kim (2011)[16] included spatial criteria in propensity score matching that was then used to evaluate a place-based crime prevention program in Seattle at the street segment level. See Walker, Winston, and Rankin (2009) for an alternative matching methodology with some spatial dimensions.

## Conclusion and Outlook

I have argued that existing spatial concepts, methods, and tools can add value to the present use of geographic data in place-based evaluations, especially by considering the implications of MAUP and spatial dependence and through the application of statistical methods to detect and control for spatial effects. In closing, I outline three important research gaps at the intersection of spatial analysis and evaluation that await to be addressed to more fully realize the potential benefits that spatial methods can have for place-based evaluations. These gaps are related to exploratory spatial data analysis, spatial modeling, and real-time learning systems.[17]

Much of the focus of exploratory spatial methods has been on spatial correlations in a cross-sectional context, rather than on quantifying differences across space for different time periods. Rey's (2001) work on LISA Markov chains is an important exception, but significant local cluster changes are measured for all time periods jointly, which does not allow for the detection of local space–time interactions. There is a need to extend existing spatial tools to explore and visualize significant improvements or declines in outcomes beyond correlations that are not benchmarked. Such tools need to identify differences over time and space and incorporate statistics to tell if these differences are significant. This approach would allow users to interactively and efficiently explore the data to see if indicators changed over time for flexible groups of subareas.

Second, there is a need for a more systematic connection between spatial econometrics and program evaluation to adjust existing spatial econometric approaches and models to better estimate benchmarked changes in relationship to program interventions and, for instance, to incorporate spatial effects with models of selection processes (Heckman 2010).

The third key research opportunity is related to current notions of rethinking evaluations by implementing a culture of "real-time learning and decision-making" (Brown 2010) that integrates evaluations with decision workflows. Such approaches provide an opportunity not only to move beyond static place-based evaluations but also to find alternatives for the typical manual process of generating static maps in evaluation. One such opportunity is to think of spatial analysis as part of real-time learning systems to monitor key indicators, which are already being tested in areas such as crime reduction or health surveillance. The idea is to feed data in near-real time to analytic models that detect patterns (such as benchmarked trends of clusters or significant differences over time); these patterns are then visualized in web-based dashboard systems on which users can manipulate inputs and outputs relatively easily. Such systems can contain the data and indicators needed by multiple organizations and programs across neighborhoods to centralize data and analysis efforts and make them comparable [see also Kramer, Parkhurst, and Vaidyanathan (2009) on shared measurement systems and RW Ventures's Dynamic Neighborhood Taxonomy (Weissbourd et al. 2009)]. An example of such an (open-source) framework is Pahle's (2012) so-called complex systems framework, which combines analytic models with scenario planning, collaboration, and decisionmaking in complex settings in a cyber framework that can use advanced computational resources to reduce model run times. The data needs of such frameworks are aided by current initiatives to make more public data accessible and integrate technology with public services (such as Government 2.0 and Code for America efforts or Chicago's big data initiative), as well as by innovative initiatives that collect data through crowdsourcing (e.g., Ushahidi).

The goal of this essay has been to inspire the use of a broader spatial perspective in the research and practice of place-based evaluation and illustrate with a few examples how spatial concepts and methods can add value in this context. After this avid endorsement of spatial analysis, it is imperative to note the importance of contextualizing such analysis within the broader questions and frameworks that evaluation research and practice have been engaged in (Kubisch et al. 2010), similarly to how quantitative impact measurement and evaluation for accountability have recently been extended to include more qualitative evaluation for learning approaches. This contextualization is especially relevant because the focus on space can reveal and account for spatial patterns, but it often does not explain the underlying mechanism (the spatial process) that generated these patterns (this is especially true when techniques such as spatial fixed effects or trend surfaces are used). In these cases, spatial patterns point to other dynamics (such as social processes) that are driving observable outcomes. Understanding the mechanisms of how and why communities change and how they can be altered to improve community residents' lives is something that goes beyond spatial analysis, but I hope to have persuaded readers of the value of an explicit spatial perspective in contributing to this larger endeavor.

## NOTES

1. See N. Verbitsky, *Associational and Causal Inference in Spatial Hierarchical Settings: Theory and Applications.* PhD dissertation, Department of Statistics, University of Michigan, Ann Arbor, MI, 2007.

2. MDRC's description of its evaluation goals of the Neighborhood Change Program (NCP) illustrate this point: "For methodological reasons, the study will not permit a formal assessment of the impacts of NCP across Chicago. Nonetheless, the research

will explore ways that a variety of social, economic, and policy factors may be working together to support community change and will explore the possibility that concentrated investments in certain places may be in alignment with positive local changes." Accessed May 2, 2012. http://www.mdrc.org/project_15_85.html.

3. *Contextual causality* has three interdependent levels of neighborhood effects: (1) spatial dynamics and higher-order structures, (2) individual selection, and (3) neighborhood composition and social mechanisms.

4. See also Cressie's (1991) taxonomy of lattice data, event point data, and geostatistical data.

5. C. Hanssen and M. Durland, "Maximizing Evaluation Impact by Maximizing Methods: Social Network Analysis Combined with Traditional Methods for Measuring Collaboration." Paper presented at the annual meeting of the American Evaluation Association, Baltimore, 2007.

6. For more comprehensive and advanced coverage, see, for instance, Haining (2003) for a general overview. Schabenberger and Gotway (2005) as well as Cressie and Wikle (2011) cover Bayesian hierarchical modeling and spatiotemporal modeling, which are relevant to neighborhood effects and assessments of change over time.

7. At Arizona State University's GeoDa Center for Geospatial Analysis and Computation.

8. Also, continuously updated lists of current spatial software development efforts are available online. For selected active (mainly open-source) programs, see http://en.wikipedia.org/wiki/List_of_spatial_analysis_software. For a longer, albeit somewhat overwhelming, list of programs with a focus on GISs, see http://www.spatialanalysisonline.com.

9. See also the OpenStreetMap-based Open Source Routing Machine, which runs fast for large datasets.

10. The software will be available for download at http://geodacenter.asu.edu/software.

11. For an overview of mostly free visualization libraries, see http://selection.datavisualization.ch/.

12. For recent reviews on new econometric developments in program evaluation, see Heckman (2010); Abbring and Heckman (2007), who include a new focus on social interactions by economists; Imbens and Wooldridge (2009); Gangl (2010); and Pearl (2009). Research applying experimental or quasi-experimental research design and statistical methods to program evaluation is well-established (Shadish et al. 2002).

13. For Verbitsky (2007), see note 1.

14. A simple univariate technique that has been applied to explore and test displacement effects (rather than model them in a multivariate context) is based on weighted displacement quotients (Bowers and Johnson 2003). Jerry Ratcliffe offers an MS Excel application to compute weighted displacement quotients based on crimes in a target area, adjacent area, and control area before and after an intervention; this application can be downloaded for free at http://www.jratcliffe.net/software/wdq-spreadsheet-calculator/.

15. J. Fagan and J. MacDonald, "Order Maintenance in Hot Spots: The Effects of Local Police Surges on Crime in New York." Paper presented at the annual ASC meeting, Washington, DC, November 15, 2011.

16. K. Kim, "Propensity Score Matching Methods for Place-Based Evaluation: Assessing the Effectiveness of the Weed and Seed Program in Seattle." Paper presented at the National Institute of Justice Crime Mapping Conference, Miami, FL, October 19–21, 2011.

17. We are working on addressing some of these research gaps at the GeoDa Center for Geospatial Analysis and Computation at Arizona State University.

## REFERENCES

Abbring, J. H., and J. J. Heckman. 2007. "Econometric Evaluation of Social Programs, Part III: Distributional Treatment Effects, Dynamic Treatment Effects, Dynamic Discrete Choice, and General Equilibrium Policy Evaluation." In *Handbook of Econometrics, Volume 6B,* edited by James J. Heckman and Edward E. Leamer (5145–303). Amsterdam, New York, and Oxford: Elsevier Science, North-Holland.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models.* London: Kluwer.

Anselin, L. 1990. "What Is Special about Spatial Data? Alternative Perspectives on Spatial Data Analysis." In *Spatial Statistics: Past, Present and Future,* edited by D. A. Griffith (63–77). Institute for Mathematical Geography Monograph Series. Ann Arbor, MI: IMaGe.

Anselin, L. 1995. "Local Indicators of Spatial Association—LISA." *Geographical Analysis* 27 (2): 93–115.

Anselin, L. 2012. "From SpaceStat to CyberGIS: Twenty Years of Spatial Data Analysis Software." *International Regional Science Review* 35 (2) April: 131–57.

Anselin, L., and D. Arribas-Bel. 2011. "Spatial Fixed Effects and Spatial Dependence." GeoDa Center for Geospatial Analysis and Computation working paper. Tempe, AZ: Arizona State University.

Anselin, L., S. Sridharan, and S. Gholston. 2007. "Using Exploratory Spatial Data Analysis to Leverage Social Indicator Databases: The Discovery of Interesting Patterns." *Social Indicators Research* 82 (2): 287–309.

Arraiz, I., D. M. Drukker, H. H. Kelejian, and I. R. Prucha. 2010. "A Spatial Cliff-Ord-Type Model with Heteroscedastic Innovations: Small and Large Sample Results." *Journal of Regional Science* 50: 592–614.

Bailey, T. C., and A. C. Gatrell. 1995. *Interactive Spatial Data Analysis.* New York: John Wiley and Sons.

Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2004. *Hierarchical Modeling and Analysis for Spatial Data.* Chapman and Hall/CRC.

Bellefontaine, T., and R. Wisener. 2011. "The Evaluation of Place-Based Approaches: Questions for Further Research." Policy Horizons Canada research paper. Ottawa, ON: Government of Canada. June.

Bloom, H. S. 2005. "Randomizing Groups to Evaluate Place-Based Programs." In *Learning More from Social Experiments: Evolving Analytic Approaches,* edited by Howard S. Bloom. New York: Russell Sage.

Bowers, K. J., and S. D. Johnson. (2003) "Measuring the Geographical Displacement and Diffusion of Benefit Effects of Crime Prevention Activity." *Journal of Quantitative Criminology,* 19(3): 275–301.

Brown, P. 2010. "Evaluating and Learning from Community Change Efforts." In *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts,* edited by A. C. Kubisch, P. Auspos, P. Brown, and T. Dewar. Washington, DC: Aspen Institute.

Clampet-Lundquist, S., and D.S. Massey. 2008. "Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment." *American Journal of Sociology* 114 (1): 107–43.

Coulton, C.J., J. Korbin, T. Chan, and M. Su. 2001. "Mapping Residents' Perceptions of Neighborhood Boundaries: A Methodological Note." *American Journal of Community Psychology* 29: 371–83.

Coulton, C., B. Theodos, and M.A. Turner. 2009. *Family Mobility and Neighborhood Change: New Evidence and Implications for Community Initiatives.* Washington, DC: Urban Institute.

Cressie, N. 1991. *Statistics for Spatial Data.* New York: Wiley.

Cressie, N., and C.K. Wikle. 2011. *Statistics for Spatio-Temporal Data.* Hoboken, NJ: Wiley.

Diggle, P.J. 2003. *Statistical Analysis of Spatial Point Patterns.* London: Oxford University Press.

Duque, J.C., L. Anselin, and S.J. Rey. 2012. "The Max-*P*-Region Problem." *Journal of Regional Science* 52 (3) August: 397–419.

Durlauf, S.N. 2004. "Neighborhood Effects." In *Handbook of Regional and Urban Economics* (4), edited by J.V. Henderson and J.F. Thisse (2173–2242). Elsevier.

Ellen, I.G., M.H. Schill, S. Susin, and A.E. Schwartz. 2001. "Building Homes, Reviving Neighborhoods: Spillovers from Subsidized Construction of Owner-Occupied Housing in New York City." *Journal of Housing Research* 122: 185–216.

Fischer, M.M., and A. Getis, eds. 2010. *Handbook of Applied Spatial Analysis. Software Tools, Methods and Applications.* Berlin: Springer Verlag.

Fotheringham, A.S., C. Brunsdon, and M. Charlton. 2002. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships.* West Sussex, United Kingdom: John Wiley and Sons.

Galster, G., K. Temkin, C. Walker, and N. Sawyer. 2004. "Measuring the Impacts of Community Development Initiatives: A New Application of the Adjusted Interrupted Time-Series Method." *Evaluation Review* 28 (6) December: 502–38.

Gangl, M. 2010. "Causal Inference in Sociological Research." *Annual Review of Sociology, 36:* 21–47.

Goering, J., and J.D. Feins. 2003. *Choosing a Better Life? Evaluating the Moving to Opportunity Social Experiment.* Washington, DC: Urban Institute Press.

Goodchild, M.F., and D.G. Janelle. 2004. "Thinking Spatially in the Social Sciences." In *Spatially Integrated Social Science,* edited by M.F. Goodchild and D.G. Janelle (3–17). Oxford, United Kingdom: Oxford University Press.

Haining, R. 2003. *Spatial Data Analysis: Theory and Practice.* Cambridge University Press.

Heckman, J. 2010. "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy." *Journal of Economic Literature* 48 (2): 356–98.

Imbens, G.W., and J.W. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.

Kelejian, H.H., and I.R. Prucha. 2007. "HAC Estimation in a Spatial Framework." *Journal of Econometrics* 140: 131–54.

Koschinsky, J. 2013. "The Case for Spatial Analysis in Evaluation to Reduce Health Inequities." *Evaluation and Program Planning.* 36(1): 172–6.

Kramer, M., M. Parkhurst, and L. Vaidyanathan. 2009. *Breakthroughs in Shared Measurement and Social Impact.* Boston: FSG Social Impact Advisors.

Kubisch, A.C., P. Auspos, P. Brown, and T. Dewar. 2010. *Voices from the Field III: Lessons and Challenges from Two Decades of Community Change Efforts.* Washington DC: Aspen Institute.

Kulldorff, M. 2010. *SaTScan v 9.0: Software for the Spatial and Space-Time Scan Statistics.* Information Management Services, Inc.

Kulldorff, M., R. Heffernan, J. Hartman, R. Assunçao, and F. Mostashari. 2005. "A Space-Time Permutation Scan Statistic for Disease Outbreak Detection." *PLoS Medicine* 2 (3): 216.

LeSage, J.P., and R.K. Pace. 2010. "Spatial Econometric Methods for Modeling Origin-Destination Flows." In *Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications,* edited by M.M. Fischer and A. Getis (409–34). Berlin Springer Verlag.

Liu, X., and J. LeSage. 2010. "Arc Mat: A MATLAB-Based Spatial Data Analysis Toolbox." *Journal of Geographical Systems* 12 (1): 69–87.

Mayne, J. 2001. "Assessing Attribution through Contribution Analysis: Using Performance Measures Sensibly." *Canadian Journal of Program Evaluation* 16 (1): 1–24.

National Research Council. 2006. *Learning to Think Spatially: GIS As a Support System in the K-12 Curriculum.* Washington, DC: National Academies Press.

Okabe, A., T. Satoh, and K. Sugihara. 2009. "A Kernel Density Estimation Method for Networks, Its Computational Method and a GIS-based Tool." *International Journal of Geographical Information Science* 23 (1): 7–32.

Okabe, A., and I. Yamada. 2001. "The k-Function Method on a Network and Its Computational Implementation in a GIS." *Geographical Analysis* 33 (3): 271–90.

Openshaw, S. 1984. "The Modifiable Areal Unit Problem." *Concepts and Techniques in Modern Geography* 38: 41.

Openshaw, S., and P.J. Taylor. 1979. "A Million or So Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem." In *Statistical Methods in the Spatial Sciences,* edited by N. Wrigley (127–44). London: Pion.

Pahle, R. 2012. "Complex Systems Framework: Integrating Analytics, Visualization and Collaboration." Working paper. Phoenix: Arizona State University. April.

Patton, M.Q. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use.* New York: Guilford Press.

Pearl, J. 2009. "Causal Inference in Statistics: An Overview." *Statistics Survey, 3:* 96–146.

Quercia, R.G., and G.C. Galster. 2000. "Threshold Effects and Neighborhood Change." *Journal of Planning Education and Research* 20:146–62.

Radil, S.M., C. Flint, and G. Tita. 2010. "Spatializing Social Networks: Using Social Network Analysis to Investigate Geographies of Gang Rivalry, Territoriality, and Violence in Los Angeles." *Annals of the Association of American Geographers* 100 (2): 307–26.

Renger, R., A. Cimetta, S. Pettygrove, and S. Rogan. 2002. "Geographic Information Systems (GIS) As an Evaluation Tool." *American Journal of Evaluation* 23 (4) December: 469–79.

Rey, S. 2010. "Spatial Empirics for Economic Growth and Convergence." *Geographical Analysis* 33 (3): 195–214.

Rey, S.J., L. Anselin, D.C. Folch, D. Arribas-Bel, M.L. Sastre-Gutierrez, and L. Interlante. 2011. "Measuring Spatial Dynamics in Metropolitan Areas." *Economic Development Quarterly* 25 (1) February: 54–64.

Sampson, R. 2012. *Great American City: Chicago and the Enduring Neighborhood Effect.* Chicago: University of Chicago Press.

Schabenberger, O., and C. Gotway. 2005. *Statistical Methods for Spatial Data Analysis.* Chapman and Hall/CRC.

Shadish, W.R., T.D. Cook, and D.T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* Boston: Houghton Mifflin.

Tobler, W.R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46: 234–40.

Verbitsky Savitz, N., and S.W. Raudenbush. 2009. "Exploiting Spatial Dependence to Improve Measurement of Neighborhood Social Processes." In *Sociological Methodology,* edited by Yu Xie. Hoboken, NJ: John Wiley & Sons.

Walker, C., F. Winston, and S. Rankin. 2009. *Assessing Community Development Program Performance: Quantitative Assessment Strategies for the LISC Sustainable Communities Initiative.* New York: Local Initiatives Support Corporation.

Waller, L.A., and C.A. Gotway. 2004. *Applied Spatial Statistics for Public Health Data.* Hoboken, NJ: Wiley.

Weisburd, D., E. Groff, and S. Yang. 2012. *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem.* Oxford: Oxford University Press.

Weisburd, D., N.A. Morris, and E. Groff. 2009. "Hotspots of Juvenile Crime: A Longitudinal Study of Arrest Incidents at Street Segments in Seattle, Washington." *Journal of Quantitative Criminology* 25 (4): 443–67.

Weiss, C.H. 1995. "Nothing As Practical As Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families." In *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts,* edited by J.P. Connell, A.C. Kubisch, L.B. Schorr, and C.H. Weiss. Washington, DC: Aspen Institute.

Weissbourd, R., R. Bodini, and M. He. 2009. "Executive Summary." In *Dynamic Neighborhoods: New Tools for Community and Economic Development.* Chicago: RW Ventures.

Yamada, I., and J.C. Thill. 2010. "Local Indicators of Network-Constrained Clusters in Spatial Patterns Represented by a Link Attribute." *Annals of the Association of American Geographers* 100 (2): 269–85.

# 8

# The Potential and the Way Forward

This book began by noting exploratory efforts in the early 1990s to assemble local administrative data from multiple sources and organize them into orderly information systems at the neighborhood level. These efforts marked a revolution at the time. Such systems moved the availability of neighborhood-level data between Censuses from almost nothing to an often fascinating something that seemed filled with possibilities. And it established a nascent field of organizations, professionals, and residents dedicated to turning data into actionable information for communities. As the previous chapters in this book have documented, the community information field did not stand still after that. The past 20 years have seen advances in the availability of neighborhood data, in the ability of individuals to manipulate them, and in their practical application.

In this final chapter, we assess where the community information field is headed now and what should be done to strengthen it. There are three sections. The first looks at the forces behind many of the changes we have described and speculates about the environment their interaction might produce in the decade ahead. We conclude that the future of this field is indeed promising. Further advances in data and technology seem inevitable, and there is momentum toward improvements in institutional arrangements and effective applications of the data. However, needed improvements in these latter two areas will not happen automatically.

Important uncertainties and barriers stand in the way. Deliberate efforts will be required if the field is to take advantage of its full potential.

Accordingly, the final two sections offer recommendations as to how current gaps and barriers might be overcome. The second section discusses what is needed to create effective data environments at the local level in this new era. The third offers ideas about the development of a national support system to encourage creating and sustaining such environments in communities across America.

## The Forces at Work and How They Might Play Out

In this section we review recent advances in key areas: institutions; data and technology; and practice and research. We close by assessing what these trends together imply for the community information field.

### Advances in Institutions

In this field, probably the most important institutional advance over the past two decades has been the development of many new local data intermediaries and the expansion of their functions. As described in chapter 2, the NNIP network has grown from 6 to 37 cities in this period; examples of how the local partners have expanded their data holdings and activities are described in chapters 3 through 5. To be sure, like most nonprofits, these organizations still struggle to survive. However, the data intermediary model has proved quite sustainable and is expanding on its own with only a small share of support from national funders. At this point, institutions in 12 additional cities working to develop similar capabilities have been in conversation with NNIP staff about joining the network.

Other local players have stepped up their activity in this field as well. First, many local governments have improved their own data capacities, making progress in staff knowledge, data collection, and program and policy applications. Fleming's essay in chapter 2 discusses this, pointing out two developments that have been particularly noteworthy: (1) the internal integration of parcel-level data from multiple departments concerned with property and housing to create enterprise geographic information systems (GIS) and (2) the implementation of data-intensive performance management programs, such as CitiStat.

Finally, as discussed in chapter 3, some cities are moving ahead with an extensive open data agenda.

Second, NNIP partners generally report there are more nongovernmental entities in their cities working with spatial data now than when they began their own operations, including other university and nonprofit research institutions and consulting firms. This change is expected because low-cost GIS capabilities and related training are now available almost ubiquitously in urban America, and more geographically identifiable data are available to the public from all sectors. As discussed in chapter 3, civic developers of data and tools are also increasing the visibility of local data use through their apps and code-a-thons.

Finally, some national organizations are also helping localities make better use of data. These include the Federal Reserve System, whose regional bank staffs have been active in promoting the development and use of community data in the past few years. In addition, long-standing organizations like the International City/County Managers Association and the National League of Cities, whose missions focus on the operation of governance at the local level, are promoting best practices among their members. Newcomers to the field from the technology sector include Code for America and DataKind, also mentioned in chapter 3.

## Advances in the Availability of Data and Technology

Recent years have brought notable progress in the willingness of government agencies to share their administrative data. The initial work of the NNIP data intermediaries to build the relationships and trust that made such sharing possible was as important as their technical work in systems building. Once a few local agencies had begun to regularly share their data with an NNIP partner (and thereby the public) the precedent was set, and it was generally easier to get other agencies to follow suit. And, year after year, the ongoing technological revolution reduced the cost of data assembly, storage, and manipulation for the intermediaries. As documented in chapter 3, these factors enabled NNIP partners to substantially expand their data holdings over the past two decades. For example, cases in chapters 5 and 6 show how integrated data on properties are now being applied productively (note, for example, the varied applications of the NEO CANDO system in Cleveland, Ohio, documented in Nelson's essay in chapter 5). Chapter 3 also discusses the growth of integrated data systems with data from multiple agencies linked at the individual level.

Relevant data available for use by NNIP partners and other local users have also markedly expanded with the federal government's public release of national files that contain consistently defined data for small areas (reviewed in chapter 3). Examples include the Home Mortgage Disclosure Act dataset and the Census Bureau's American Community Survey. Commercial firms have published a number of new nationwide datasets with address-level data as well. Examples include property sales data from firms such as Experian and data on business establishments from firms such as InfoUSA. Chapter 3 also notes that the open data movement will result in many more datasets at the local, state, and federal levels being made available to the public online; see also Pettit, Hendey, and Kingsley (2014) and Tauberer (2011). The impact will be particularly dramatic in communities without local data intermediaries.

Finally, we have to recognize the possibility that other datasets now being talked about in the realm of "big data" might be brought to bear on issues relevant to local governance. For example, data from sensors may produce neighborhood-level noise or air quality indicators for the first time. Data from Twitter may be able to be translated into subcity-level indicators of neighborhood social networks and community satisfaction.

In addition to expanded availability of data, the years since 1990 have seen a number of technological advances that facilitate their use. The number of data dissemination and visualization platforms is growing rapidly. Some sites simply provide more efficient access to data, like government open data portals or the National Historical Geographic Information System. Other websites include visualization tools, such as PolicyMap and Esri's Community Analyst.

Federal and other agencies also provide decision-support tools developed for specific purposes. These tools bring together data from different sources and manipulate them in preestablished, automated routines to produce displays that directly support specific decisionmaking tasks. One example is the US Department of Housing and Urban Development's eCon Planning Suite (described in chapter 3), which provides an online template and data tools to help the department's grantees complete their required consolidated plan (US Department of Housing and Urban Development 2012). Such tools can help practitioners navigate the avalanche of data that is currently accessible, presenting the data most relevant to the task at hand in a form that helps them assess alternative courses of action efficiently.

## Advances in Practice and Research

As documented in earlier chapters of this book, many individual success stories and promising attempts illustrate how community information is being used to advance both practice in local governance and community work and the research base needed to understand neighborhood conditions and programs. And momentum seems to be growing. Still, further advances in data use are not likely to be as automatic as they are with respect to data availability, and there remains plenty of room for improvement. We believe that the application of the data is the area in which the field's leaders need to concentrate their efforts. The ultimate goal should be to reach a state in which data are routinely being taken advantage of and data-driven decisionmaking is fully incorporated into programs and policies to attain societal goals at the local level in America. Achieving this goal would include enabling all constituencies, especially those with limited access to power, to use data effectively to voice their interests in local decisions.

Chapter 4 introduces a short list of basic applications of neighborhood-level data and discusses how the literature suggests each might function in an advanced form. Our conclusions regarding the current state of practice and research are as follows:

- *Situation analysis.* Even with the growth of community indicators projects described in chapter 6, the process of regularly reviewing a broad array of indicators on multiple topics at the neighborhood and city levels has not been fully institutionalized in most places. These efforts need to work harder to incorporate analysis of trends in multiple phenomena in relation to a well-structured logic model, with an eye to gaining a sense of how the underlying forces are likely to interact to affect outcomes over the years ahead.
- *Policy analysis and planning.* Many local agencies and initiatives now use data in some ways in these areas. However, here too there remains ample room to do better. Few now consistently apply the most advanced practices in decisionmaking, such as orderly planning processes in which analytic models (or simpler techniques) are used to generate and compare explicit estimates of outcomes (measured costs and benefits) implied by alternative programmatic scenarios.
- *Performance management.* As noted earlier, funders and the public are increasingly pressing local agencies to use data for continuous

program improvement. Efforts to regularly quantify and report on progress in relation to goals in a meaningful way are now underway in a sizable number of settings. However, performance management is still at a fairly early stage. Many relevant local institutions have not yet put serious performance management systems, in particular systems that take advantage of neighborhood-level data, into operation.

- *Education and engagement.* Examples reviewed in chapters 5 and 6 suggest that local practitioners and advocates have developed considerable expertise in using data for education and engagement. Encouraging local constituencies to bring a critical perspective to what the data show will lead to new insights into neighborhood issues and better-tailored responses. More creativity in these practices is certainly warranted. Our judgment, however, is that progress on this front will also rely on improvements in the number and quality of the applications noted above. Compelling analysis of local data that contributes to understanding and solving real problems is needed to motivate new coalitions and drive policy and program changes.
- *Neighborhood research.* As chapter 7 demonstrates, progress in data availability and technology has enabled many advances in analytic methods over the past two decades. A number of researchers have put forward new concepts and explored and tested them with the new data in ways that would not have been possible in the 1990s. However, we see considerable untapped potential. In particular, we need a more robust evidence base for understanding neighborhood dynamics and policy interventions. Only with focused attention will we be able to tackle the many unanswered questions that remain about how to improve neighborhoods and support better outcomes for residents.

## Outlook: Status and Challenges

Considering all the topics above, what is likely to happen if current trends continue? As noted at the start of this chapter, there seems little doubt that the availability of relevant data will expand. Unrelated advances in basic technology, in particular, have substantial momentum. The difficulty and cost associated with data assembly, storage, and manipulation will continue to decline.

We also see a positive sign in the increasing pressure from funders and the public at large for data-driven decisionmaking by all the institutions involved in local governance. In the 1990s, government and civic leaders had little direct exposure to how data might influence collective behavior locally. Since then, as discussed in chapter 4, there have been more examples of the effective use of data in performance management. We think it is the coupling of these examples (showing that it is possible to make a difference) with the interest in making real progress in societal issues that have fueled increased pressure for accountability. This pressure is clearly being applied to a broad spectrum of nonprofits [with calls for transformations in culture; see Morino (2011)] as well as to government agencies. And although the focus now is on performance management, we think it will come to be understood that more effective data-driven situation analysis and policy analysis and planning are as important to outcomes as the persistent tracking of results.

We also recognize the major advances in policy applications and methodology are well beyond what might have been expected in the early 1990s. It is in these areas, however, that we believe that the most additional progress needs to be made. It is not assured that progress in this field will be as rapid as it could or should be, considering the urgency of the issues at hand. Major barriers remain to be overcome.

First, although no authoritative inventory of status in communities across the nation exists, conversations with a sizable number of local contacts convince us that civic leaders in many places are not yet well-informed about the benefits of an enriched local data environment. Even though interest is growing, NNIP partner institutions operate in only three dozen US metropolitan centers, a small fraction of the total.[1] The cities most well-known for innovative open data policies and practices are the largest cities. Relevant work is being done in other places, but the fragmentary information we have obtained suggests that although more local institutions are working with data, this work often lacks coherence and momentum.

Second, even in the places that have seen the most promising work with cross-sector administrative and geographic data, the future is not clear. It is not clear that the most advanced practices (e.g., extensive property data systems and integrated data systems with linked information on individuals and families) will be sustainable over the long run. They involve difficult technical work in a real and risky political environment. We cannot be certain they will be brought to scale, sustained, and replicated. Many agency personnel still remain nervous about sharing data, which slows the release of data to other collaborating agencies, let alone to the public via data intermediaries or open data platforms.

As implied above, the level of data use lags behind the level of data production. Given today's skills, habits, and cultures, much work remains before it will be possible to characterize all relevant local institutions as being data driven. Too many key players still see work with data as an extra task that takes time away from their main job, rather than something that is an essential component that saves time and improves the quality of their work. It is not yet built into their daily work routines. In some cities where agency datasets have been released to the public directly on open data platforms, there is a concern about underutilization at this point (Pettit et al. 2014). Our research community needs to intensify its work in mining the available data to generate scientific hypotheses and studies. Although progress is being made, more needs to be done to strengthen the institutions, access to data and technology, and ability to put the data to use for practice and research.

However, although there is uncertainty, we see the glass as much more than half full. Even though the pace feels painfully slow at times, local data capacity seems to be taking forward steps everywhere. Looking back over the past 20 years, more administrative agencies are sharing data, and productive applications are advancing, not retreating. We have noted the increased pressure for accountability of public and nonprofit entities via data in the past few years and definitely see an increase in the buzz about data use at the local level generally. In fact, we believe that local governance, community decisionmaking, and urban research in America could be at the edge of a major transformation as data availability continues to expand and innovations in training and tools make the data much easier to use. The final two sections of this chapter contain our recommendations for making this transformation a reality.

## Creating an Effective Local Data Environment

National researchers and policy makers can take advantage of the types of data examined in this book to better understand conditions and trends in America's communities, but by far the most important applications occur at the local level. Examples involve the use of small-area data to improve individual neighborhoods (chapter 5) and to strategically address city- and metropolitanwide policy issues where spatial variations make a difference to outcomes (chapter 6). Although most metropolitan areas do not have fully coherent governance structures, most of those areas have identifiable coalitions of civic leaders (the regimes discussed in chapter 2). In all locations where local data intermediaries have been established, the local leadership regime has been actively involved, most often taking the lead in making it happen. These are the people whose support needs to be mobilized to create effective local data environments where they do not exist or to strengthen them where they are already in process. What should we be asking them to do?

## Assign Primary Responsibility and Strengthen Support for Local Data Use

Our first recommendation in this area is that civic leaders assign one or more institutions clear responsibility to play the leading role in building an effective data environment. This responsibility starts with a commitment to the key functions outlined in chapter 2: assembling and disseminating neighborhood-level data from multiple administrative agencies and facilitating practical applications of the data to achieve impact. The experience of the last two decades suggests that the benefits of local data are only likely to be achieved when these basic responsibilities have been clearly assigned to some institution (or a collaboration of institutions), and these institutions are supported in that role over the long term.

For metropolitan areas, the local data intermediary model presented in chapter 2 probably offers the most promising approach. For small cities, towns, and rural areas, the local environment may not be able to financially support a dedicated data intermediary. In addition, the term *neighborhood* may take on a different meaning. In these cases, data support may be provided by a different institutional arrangement. Possibilities include a statewide entity, such as the one envisioned by the Utah Community Data Project,[2] or a center covering several smaller towns, like the SouthCoast Urban Indicators Project, which supports Fall River and New Bedford, Massachusetts.

It might be suspected that with more data available in cities where the open data movement is moving forward, the importance of an intermediary might diminish. In contrast, a recent review (Kingsley, Pettit and Hendey, 2013), found that demand for intermediary services is increasing, not contracting. Datasets released on open data platforms are often released in a raw form, which can markedly simplify the data assembly function for intermediaries. However, a high level of technical skill is required to use them directly, and datasets are typically released one by one. In these situations, intermediary services are still needed to put the data in a form that will make them easier for others to use and then to combine disparate datasets so they can be used together.

Who should pay for the operation of the new local data intermediaries? Chapter 2 explains that funding for the NNIP partners since 1995 has come mostly from local sources (aided by the explicit or implicit endorsement of local civic leadership coalitions). Most partners earn a good share of their keep by providing specific studies and other data-related services under contract to a variety of clients, public and private. But all of them also receive some general support funding from local philanthropies. Their basic intermediary services can be regarded as a part of civic infrastructure and, as such, warrant local subsidy. We believe that the spread and strengthening of local data intermediaries (or other responsible entities) should be mostly funded in a similar manner. However, in our recommendations about a national support system in the last section of this chapter, we also propose that some federal and state funds be provided to help in the start-up phase and to incentivize innovation thereafter.

## Broaden and Coordinate Participation

When NNIP began it was expected that civic leaders would normally assign the data intermediary role to one local institution that would be responsible for the basic functions of building and operating an ongoing neighborhood-level information system and ensuring its active use in the public interest. Since then, the network has recognized the need to broaden its horizons with respect to both functions and participants; see "Using Data More Broadly to Strengthen Civic Life and Governance" in chapter 2. We noted earlier in this chapter that in most localities, a much broader range of local institutions is likely to be involved in data today. Our conclusion is that this broadening of participation should be encouraged, but it also needs to be coordinated. Stories of duplication, as well as gaps, are now more frequently heard.

Wascalus and Matson note in their essay at the end of chapter 2 that one critical component of the Twin Cities' strong community information ecosystems was "forging the interagency and professional connections needed both to exchange ideas and to advance the overall data-sharing environment." We believe that in most places, new efforts will be needed to promote coherence, reduce fragmentation, and provide forums for joint deliberations on how best to take advantage of new opportunities.

Civic leaders should first examine their local data environments holistically, both the individual institutions and the current forums for communication. With this understanding, they can determine locally appropriate institutional arrangements to forge better connections among the players and conduct joint activities. Different institutional models will be appropriate for this in different places. Wascalus and Matson's essay describes various regional advisory councils and other topical forums in which staff from information-related organizations in the Twin Cities share plans and ideas. The interactions are embedded in an environment in which the local institutions have been working collaboratively for years to expand data availability and productive applications. In contrast to this organic model, we see advantages in establishing a formal community information network that all relevant local players would be encouraged to join.[3] The network could be coordinated by an existing institution and would increase the chances that communication and plans would occur despite the competing priorities of individual organizations. In addition, connections between organizations would be cemented by the network structure and would not rely solely on one-on-one relationships between staff members who may move on.

Whatever the form, the coordinating group(s) could serve the following functions:

- Convene regular meetings of community information network stakeholders to share plans, discuss ideas about new opportunities, and consider ways to better coordinate the work;
- Promote periodic events to bring users of the data together to highlight successes and challenges and to build users' and providers' skills and leadership; and
- Use various strategies to communicate the importance of community information to government officials, funders, and the broader public.

Two other activities the community information network might support could be particularly important in enhancing capacity. The first would be to offer training to local stakeholders on how to use data to support the work of their own institutions. The training would spread best practices in all types of the applications discussed in chapter 4 (situation analysis, policy analysis and planning, performance management, education and engagement, and research on neighborhood dynamics and effects). A basic training could give an overview on the sources of data that are available locally, with a focus on applying the data to address real problems in various fields. A more advanced training example is the Certificate in Program Design and Development program focused on performance management offered by the Institute for Urban Policy Research at the University of Texas at Dallas.

A second group of new activities includes efforts to better connect data intermediaries and other actors in the community information field with efforts to advance open data locally. NNIP nationally is already working with leaders of the open data movement along these lines and has developed a concept paper explaining how the two movements can work together (Pettit et al. 2014).[4] In addition to building stronger relationships with open data advocates and civic developers, community information practitioners can assist the open data movement by

- *Promoting the open data approach both locally and in larger forums.* Promotion entails explaining open data processes and benefits to civic leaders and practitioners and participating with open data leaders in the development and implementation of a local open data agenda.
- *Helping application developers target their work productively.* This assistance entails informing developers about local issues and connecting them to other groups, particularly groups representing the interests of low-income neighborhoods. In this way, the applications can be developed collaboratively and better respond to community needs.
- *Facilitating more active and productive use of released datasets.* Such facilitation includes transforming open data releases to make them easier to use and integrate with other available datasets. It also includes explaining and promoting released datasets to local stakeholders to encourage greater utilization.

## A National Support System

Creating effective local data environments is the goal, and much of what is needed to accomplish that goal can be locally self-supporting. However, progress is likely to be unacceptably slow unless some outside motivation and support are provided. In this section we recommend a program with seven elements that together would constitute a national support system to accelerate effective local data development and use [adapted from ideas discussed in Kingsley et al. (2013)]:

1. Inform local civic leaders about the potentials of a more effective data environment.
2. Develop and connect networks to document and disseminate local best practices and provide technical assistance and training to local data intermediaries.
3. Expand access to data from governments and private firms.
4. Expand data transformation and visualization tools to facilitate local data use.
5. Develop capacity to educate future policymakers and practitioners in the methods of data-driven decisionmaking.
6. Establish a network to strengthen neighborhood research.
7. Provide funding to catalyze local institutional development and incentivize innovation.

Before reviewing the recommendations in each area, we explore how a group of institutions might collaborate to begin to implement them. In most successful local data environments today, progress has occurred primarily under nonprofit-sector leaders who are able to build strong collaborative and ongoing relationships with government agencies. We think the development of a national support system for community information is likely to evolve in the same way.

No single institution is an obvious starting place. Rather, we believe that a number of national groups with interest in this field should work together to form a coalition to carry this work forward. NNIP is already carrying out aspects of the field-building work that is required (the basic functions of the NNIP network are summarized in chapter 2), but it now serves only a segment of the broader market. Although NNIP might convene and help to launch such a coalition, many more organizations should be involved, especially those whose missions focus on local governance and community improvement.

As noted in chapter 2, many national organizations already support strengthening local data capacity in various ways and would be candidates to participate. This includes groups working in particular issue areas as well as those focused on broader social justice agendas. Representatives of national foundations that have funded the strengthening of local data capacity in the past and relevant federal agencies should also be involved from the outset.

Finally, national networks of local actors could help reach out to their memberships, including those that represent local civic funders, such as community foundations and United Ways; local governments, such as the National League of Cities and the International City/County Management Association; and local nonprofits, such as LISC and Enterprise. As noted in chapter 3, proactive efforts are needed to make sure that the increased access to data and exciting new technology contribute to inclusive community improvement. Thus, we believe that the coalition's mission statement should prioritize using data to address the problems of low-income neighborhoods and do so in a way that emphasizes the engagement of residents and neighborhood organizations. Our suggestions about the program elements are as follows:

1. Inform local civic leaders about the potentials of a more effective data environment.

The coalition would need to develop the capacity to inform a broad array of entities by making presentations on strengthening the local data environment at national conferences of the relevant institutions and writing articles, blogs, and guides on the topic for membership publications and websites.

The groups that represent local civic funders would communicate with their local affiliates, share relevant materials, and encourage them to participate in coalition-sponsored activities that illustrate the role of local data intermediaries and the benefits of an enhanced local data environment.

2. Develop and connect networks to document and disseminate local best practices and provide technical assistance and training to local data intermediaries.

As explained in chapter 2, these are activities that NNIP does now, albeit on a much smaller scale than is called for to achieve the goal set forth at the beginning of this section. There are lessons here for the coalition as it attempts to build mechanisms to support improved practice in a much broader range of localities, but simply spreading the NNIP model to other places, per se, may not be the best way to increase scale. For example, a group of separate but related networks (perhaps subdivided by region or size of city) might make more sense than one unitary network structure. There is not enough information available now to support a specific recommendation on how these supports can be most effectively delivered at the broader scale we have proposed. The coalition should consider the design for an effective peer network with these capacities once it has reviewed and cataloged existing networks and resources.

3. Expand access to data from governments and private firms.

This area entails efforts by the coalition to encourage all levels of government and some private firms to release their existing nonconfidential datasets to the public in usable formats. In all aspects of this effort, the coalition should work in partnership with the open data movement discussed in chapter 3 and earlier in this chapter (Pettit et al. 2014).

Without devoting many of its own resources to direct advocacy, the coalition could make an important contribution in this area simply by serving as a central scorekeeper, annually documenting and assessing progress with data releases and suggesting priorities for advocacy by others based on that assessment. The situation differs at different levels:

- *Federal government.* Chapter 3 notes that the number of national datasets with useful information for local analysis has grown rapidly in recent years. The federal government has taken explicit positions favoring these directions [see Orzag (2009) and White House (2009)], and considerable progress has been made under data.gov. In addition to promoting new releases, however, there is a need for advocacy to preserve current national data collection and dissemination efforts that may be threatened by funding cuts. The coalition's national scorekeeping function could be helpful in suggesting priorities for preservation as well as those for new releases.
- *State governments.* To date, reviews of open data at the state level have focused on legislative or budget data; see OpenStates.org as one example. More information is needed about other data, such as education, health, housing, and public assistance data. Additional research could document how the states are making progress opening up a broad array of data, including data that can be disaggregated for census tracts, cities, and counties.
- *Local governments.* One recommendation here is for the coalition to incorporate education about, and advocacy for, open data as an integral part of its information campaign to encourage broader

development of local data intermediaries and related capacities. The second section of this chapter discusses important roles local data intermediaries can play, not only in encouraging local momentum for open data releases, but also in making the data usable to a wider range of local stakeholders and in assuring they are generally applied in ways that yield higher payoffs.

- *Private firms.* Despite the growing recognition of the benefits of open data, some city and county governments continue either to sell public records files directly or license the data to private firms. This practice results in the data being prohibitively expensive for local users. The public has already paid for the original creation of the records within government, and on principle should not have to pay a second time to access them. To make the financial case, the coalition should mount evidence to demonstrate that the long-term and widespread benefits that the city can gain by having public and private decisions informed by open data will overshadow the limited and short-term income from selling the files. The coalition, federal agencies, and national networks would reinforce this message and require that nonconfidential data development under their funding must be open to the public.

  In addition, this coalition could explore the access of proprietary data for neighborhood-related research. Private firms collect many potential data, such as job openings, home sales or rental postings, and grocery store purchases, which would reveal new insights into neighborhood conditions. The coalition could help reach out to key companies, develop model data-sharing agreements, and facilitate the publishing of aggregate or lagged data.

4. Expand data transformation, visualization, and tools to facilitate local data use.

The first section of this chapter discusses a number of technical advances that facilitate use of the rapidly growing amount of data. These advances include the transformation of large and complex datasets to make them more usable to new visualization platforms; new decision-support tools; and the many new applications civic developers have created to make data directly usable to individuals.

In our judgment, considerable room remains for further development in all these areas. More ambitious decision-support tools, for example, are needed (such as simulation models that can more reliably predict changes in neighborhood conditions under varying assumptions about economic conditions and program actions).[5]

However, given the capacity of the players and the competition that now exists in this environment, we believe these activities are likely to move forward on their own without much outside support. The only role we recommend for the coalition is to monitor advances in a coherent way and wrap the results into its ongoing campaign to keep local practitioners and decisionmakers well-informed of the state of the art in this field. For these audiences, extra efforts need to be made to consider and elaborate the practical implications of such advances for local work.

5. Develop capacity to educate future policymakers and practitioners in the methods of data-driven decisionmaking.

Universities and community colleges that train people for careers in local policy and program development (primarily schools of public management and urban planning) already have courses in GIS and quantitative methods. However, approaches to data-driven decisionmaking such as those emerging in some NNIP settings are typically too new to have found their way into those curricula.

These approaches go far beyond basic statistical concepts and techniques. They are more rooted in policy analysis and explain how analysts and planners set the stage for strategic decisionmaking in complex institutional settings. They cover how today's richer array of neighborhood indicators can support more nuanced and realistic situation analysis (fact-based assessments of strengths, weaknesses, opportunities, and threats) than was possible in the past. They also cover more effective methods of quantitatively weighing alternative courses of action to address different community situations and making choices based on those analyses. Courses covering this content should be developed for traditional university settings and also as professional development modules and stand-alone online classes.

The coalition should mount an effort to document advances in these areas. Documentation and communication should occur at varying levels, ranging from vignette-length descriptions of innovative practices to detailed how-to guides. In-depth materials could be adapted for use in university curricula. The coalition should expect to devote considerable energy to communicating about these resources with relevant university faculty and their associations.

6. Establish a network to strengthen neighborhood research.

Once the coalition has been established and is working with universities to strengthen their curricula in the practice side of this field, it should be in a good position to propose and help establish a new network to strengthen neighborhood research. A network is needed that fosters methodological development that transcends substantive areas and promotes exchanges between individuals in academia, policy settings, local data intermediaries, foundations, and government agencies that study, operate, and invest in neighborhood programming.

As noted in chapter 7, those who have contributed to this work to date come from many institutional settings and academic disciplines. There is a need for new mechanisms to facilitate communication among them. As a first step, the coalition should identify the scientific societies that already have interest groups or forums on community-based research that could be supported and connected by a developing network. It then should meet early with the leading scholars who have made key contributions to date to discuss possible mission statements, organizational structures, and priority activities. These activities could mobilize disparate efforts to address the major areas in need of development identified in chapter 7:

- Developing more neighborhood concepts with uniform definitions and standardized measures and encouraging their widespread use;
- Taking advantage of potential for more fine-grained and dynamic spatial analysis;
- Amplifying and extending our ability to address the counterfactual problem in order to greatly strengthen evaluations of practices and policies; and
- Advancing the analytic methods needed to provide understanding of the dynamic processes that shape neighborhoods, thereby building the capacity to model the process of neighborhood change holistically.

7. Provide funding to catalyze local institutional development and incentivize innovation.

The coalition's information campaign should be designed to encourage the federal and state governments to support building local data intermediary capacity—support that would add to and incentivize expanded investment by local and national philanthropies. At present, most communities do not have these capabilities at all. In the places that do have them, more resources are needed to assure sustainability and to allow intermediaries to pursue new ideas for creative, cost-saving applications that could serve as valuable models for others.

One way this can be done is to make sound local data work a criterion for awards of competitive federal grants and to assure that those grants provide sufficient funding for such work. Other ways federal agencies could promote this field would be to include descriptions of innovative applications by local data intermediaries in their newsletters and other publications, recognize such work in their awards programs, and build it into their technical assistance programs.

In addition, broader ongoing support should be considered through the Community Development Block Grant program and the Social Innovation Fund. Explicit support for data-driven decisionmaking would seem not only consistent with, but also important to, the goals of both.

Finally, we recommend that the federal government reinstitute a program of matching grants for technical innovations, similar to the Technology Opportunities Program, which was administered by the National Telecommunications and Information Administration of the US Department of Commerce from 1994 to 2004. As discussed in chapter 3, Technology Opportunities Program funding was responsible for several of the most important advances in this field during that period. It encouraged local NNIP partners and others to reach out to create bolder changes to current practice than they otherwise would have done. These changes (e.g., web-based interactive mapping) accelerated the progress of the field as a whole.

## Conclusions

This chapter discusses the forces underlying many advances in the production and use of community information that have been made over the past two decades. Our overall conclusion is strongly positive. The future of this field appears promising, enough so that we suggest it could be the basis for a transformation in the quality of local governance, community decisionmaking, and urban research in America. We also suggest, however, that this potential will not be achieved automatically. Special efforts are needed to take advantage of the opportunities that exist.

The challenges differ in various aspects of the work. Substantial momentum already exists on the technical side. Further declines in the difficulty and costs of assembling, storing, and manipulating data seem almost assured, without new outside prompting. These declines will make sharing data easier, and the experiences reviewed in this book suggest the resistance of government data holders to sharing data is generally breaking down. More prompting is warranted here: for example, via the continued efforts of the open data movement. But overall, more and more data relevant and useful for community purposes will become available over the coming years.

We argue, however, that more effort is required to change current habits and practices pertaining to the use of such data. Many of the institutions that are key to improving conditions in low-income communities are not yet data driven, but we are beginning to see that the trend is in the right direction. The availability of more training and easy-to-use tools, as well as more data, should in themselves motivate more productive applications. However, the pace of accomplishment in these areas may be slow overall unless outside pressure and support expands significantly.

Many of the applications we have reviewed in this book do indeed reflect advanced practice. These examples include sophisticated policy applications few would have expected possible 20 years earlier. The work we have described with extensive property data systems and integrated data systems (about individuals and families) is particularly noteworthy and promising. Also impressive are the institutional capacities that have been developed in a number of cities to encourage, improve, and sustain effective data use in local policy and community improvement (particularly by the institutions that are a part of NNIP). The problem, though, is that these examples have occurred in only a few places. Only a fraction of cities, including the NNIP partner cities and a few others with community-engaged research centers or unique government offices that take on community data applications, have the capacity to do this work. The challenge then is not only to develop more advanced applications, but to tackle the more sizable task of encouraging the development of similar capacities in many more places.

How can these goals best be achieved? What counts most is what happens locally, so we first offered recommendations to the civic leaders in major urban areas. At this level, we believe the priority needs to be on building and sustaining institutional capacity. Civic leaders should carefully select one or a small group of trusted local institutions with the needed technical capacity and then charge it with developing recurrently updated data systems, applications, and processes necessary to encourage productive use of the data. The civic leaders must begin with the understanding that their goal is to build capacity that will be permanent. The selected institutions must regularly prove their worth to continue to receive support. NNIP experience offers a variety of models as to how this goal can be accomplished in differing civic environments.

Although accomplishment at the local level is the ultimate goal, we recognize that the needed improvements will also require the development of a national support system for this field. We do not offer a precise plan for how this system should be built, but suggest that it should be developed by a coalition of representatives of the federal government, national nonprofit organizations, and national networks that have a strong interest in effective local governance. Priority for using data to address the problems of low-income neighborhoods in an inclusive manner should be a part of its mission statement. The support system would provide (or arrange for) services in a number of areas based on the coalition's assessment of priorities.

Clearly, more effective use of data in local decisionmaking will not be sufficient to fully address poverty and the other major social problems of our time; such results will surely depend on fundamental changes to socioeconomic structures and policies. Nonetheless, the public and policymakers alike dramatically underestimate the scope and importance of local-level governance decisions to national well-being. Effective use of data can markedly improve the quality of those decisions. There may be few investments that offer a higher payoff than resources devoted to expanding data capacity in local communities throughout America.

## NOTES

1. It should be noted, however, that NNIP partners are typically located in larger metropolitan areas, so their influence may be greater than their fraction of the number of metro areas might seem to imply. Metropolitan areas with NNIP partners in 2012 had a total population of 119 million (2010 census), more than one-third of the total US population.

2. For additional information on the Utah Community Data Project, see http://www.ucdp.utah.edu/.

3. The need for such coordination and possible means of achieving it are discussed in a recent review of the overall data environment in Chicago (Pettit and Kingsley 2013).

4. In addition to this paper, NNIP has also published several brief case studies on how NNIP partners are already working in support of open data goals in their cities.

5. For additional examples, see Tatian (2012) and Kingsley and Pettit (2008).

## REFERENCES

Kingsley, G. Thomas, and Kathryn L. S. Pettit. 2008. Data and Decisions: Parcel Level Information Changing the Way Business Gets Done. Washington, DC: Brookings Institution.

Kingsley, G. Thomas, Kathryn L. S. Pettit, and Leah Hendey. 2013. *Strengthening Local Capacity for Data-Driven Decisionmaking.* Washington, DC: Urban Institute.

Morino, Mario. 2011. *Leap of Reason: Managing to Outcomes in an Era of Scarcity.* Washington, DC: Venture Philanthropy Partners.

Orzag, Peter R. 2009. "Open Government Directive: Memorandum for the Heads of Executive Departments and Agencies." Washington, DC: Office of Management and Budget.

Pettit, Kathryn L. S., Leah Hendey, and G. Thomas Kingsley. 2014. *NNIP and Open Data.* Washington, DC: Urban Institute.

Pettit, Kathryn L. S., and G. Thomas Kingsley. 2013. *An Assessment of the Community Information Infrastructure in the Chicago Metropolitan Area.* Washington, DC: Urban Institute.

Tatian, Peter A., G. Thomas Kingsley, Joe Parilla, and Rolf Pendall. 2012. Building Successful Neighborhoods. Paper prepared for the What Works Collaborative. Washington, DC: The Urban Institute.

Tauberer, Joshua. 2011. *Open Data Is Civic Capital: Best Practices for Open Government Data.* Accessed October 25, 2011. http://razor.occams.info/pubdocs/opendatacivic-capital.html.

US Department of Housing and Urban Development. 2012. *The eCon Planning Suite: A Desk Guide for Using IDIS to Prepared the Consolidated Plan, Annual Action Plan and CAPER/PER.* Washington, DC: Office of Community Planning and Development, US Department of Housing and Urban Development.

White House. 2009. *Open Government: A Progress Report to the American People.* Washington, DC: White House.

# About the Authors

**G. Thomas Kingsley,** a senior fellow at the Urban Institute, specializes in housing, urban policy, and governance issues. He served for over a decade as director of the Institute's Center for Public Finance and Housing and for 17 years as director and codirector of the National Neighborhood Indicators Partnership, an initiative to further the development of advanced data systems for policy analysis and community building in US cities. Since 2000, Mr. Kingsley's research topics have included analyzing neighborhood patterns and impacts of the foreclosure crisis; assessing lessons from HUD's HOPE VI program for urban policy and the future of public housing; providing analytic support to the Annie E. Casey Foundation's Making Connections Initiative; and analyzing the patterns and effects of concentrated poverty in America's urban areas. In the 1990s, Mr. Kingsley was codirector of the Ford Foundation–sponsored Urban Opportunity Program, which produced four books on the status of urban policy issues in America. Before that, he was the director of several major policy research programs, including testing the market effects of housing allowance programs (1974–80, for the HUD-sponsored Housing Assistance Supply Experiment); analyzing the structure and potentials of metropolitan Cleveland's economy (1980–82, for the Cleveland Foundation); preparing a national urban development strategy for Indonesia (1982–85, for the United Nations); and helping the Czech and Slovak Republics design and implement policy reforms in housing and municipal infrastructure (1991–95, for USAID). Mr. Kingsley previously served as director of the Rand Corporation's Housing and Urban Policy Program and as assistant administrator for the New York City Housing and Development Administration. He has taught on the faculties of the graduate urban planning programs at the University of California, Berkeley, and the University of Southern California.

**Claudia J. Coulton** is distinguished university professor, Jack, Joseph and Morton Mandel School of Applied Social Sciences, Case Western Reserve University, where she is also the Lillian F. Harris professor of urban research and social change. Her research focuses on approaches to community change and the challenges facing distressed urban neighborhoods. She is the founding director of the Center on Urban Poverty and Community Development and oversees a multidisciplinary team working to better understand the systemic forces that produce distressed neighborhoods and what individuals, organizations, and policies can do to reverse these conditions. Under her leadership, the center has built a model of how to provide data for community initiatives, including a dynamic neighborhood indicators system, a parcel-based property-tracking system, and a longitudinal multiagency integrated data system for children. Coulton is a founding partner of the National Neighborhood Indicators Partnership and has served as research advisor to many community change programs, including the Aspen Institute's Roundtable on Comprehensive Community Initiatives, the Annie E. Casey Foundation's Making Connections initiative, and the Invest in Children program. Her current studies focus on processes of residential mobility and neighborhood identity in poor communities, the effects of the foreclosure crisis on families and neighborhoods, and the impact of the built and social environments on outcomes for families and children. She is the author of numerous scientific publications and policy reports. Coulton received a BA in sociology from Ohio Wesleyan University, an MSW in social work from Ohio State University, and a PhD in social welfare from Case Western Reserve University.

**Kathryn L. S. Pettit** is a senior research associate in the Metropolitan Housing and Communities Policy Center at the Urban Institute. Her work focuses on measuring and understanding neighborhood change particularly in low-income neighborhoods. She is a recognized expert on using local and national data systems for housing and urban development research and program development. She directs the National

Neighborhood Indicators Partnership, a collaborative effort by the Urban Institute and local partners to further the development and use of neighborhood-level data systems in local policymaking and community building. Other recent work includes a study on the role of open data at the local level and an assessment of the US Department of Housing and Urban Development Choice Neighborhoods Initiative. Past projects include serving as coauthor for a seven-year series on housing markets in the Washington, DC, area; producing data and guides for the website DataPlace; and contributing to several indicators-related projects and neighborhood analyses for various federal agencies. Pettit has a master's degree in public policy and a BS in international relations from Georgetown University.

# About the Essayists

**Riccardo Bodini** received his undergraduate degree in communication studies from the University of Bologna (Italy) and a master's degree in public policy from the University of Chicago. After graduating, Bodini worked for seven years at RW Ventures LLC, an urban economic development firm in Chicago, where he collaborated with businesses, foundations, think tanks, government, and community-based organizations on a variety of applied research projects. In January 2010 he returned to Italy to join the European Research Institute on Cooperative and Social Enterprises (Euricse), where he applies his experience in research and project management to the development of new research and policy projects in the area of cooperative and social enterprise studies. Starting in 2012, he also took on the role of coordinator for Euricse, managing the day-to-day activities of the institute.

**Hedy N. Chang** directs Attendance Works, a national- and state-level initiative aimed at advancing student success by addressing chronic absence. A skilled presenter, facilitator, researcher, and writer, she coauthored the seminal report *Present, Engaged, and Accounted For: The Critical Importance of Addressing Chronic Absence in the Early Grades,* as well as numerous articles about student attendance. Deeply committed to promoting two-generation solutions to achieving a more just and equitable society, Chang has spent more than two decades working in the fields of family support, family economic success, education, and child development. She served as a senior program officer at the Evelyn and Walter Haas Jr. Fund and as codirector of California Tomorrow, a nonprofit committed to drawing strength from cultural, linguistic, and racial diversity. In February 2013, Chang was named by the White House as a Champion of Change for her commitment to furthering African American education.

**Robert J. Chaskin** is associate professor and deputy dean for strategic initiatives at the University of Chicago's School of Social Service Administration and an affiliated scholar at Chapin Hall at the University of Chicago. His work focuses on the conceptual foundations and principal strategies of contemporary community intervention in the context of urban poverty. He has written widely on the topics of neighborhood intervention, community capacity building, and the dynamics of participatory planning and neighborhood governance. His most recent publications include the edited volume *Youth Gangs and Community Intervention* (Columbia University Press 2010) and various journal articles based on his current research on policy efforts to transform public housing neighborhoods through mixed-income development schemes. He holds an AM in anthropology and a PhD in sociology from the University of Chicago.

**Cory Fleming** oversees the International City/County Management Association (ICMA) 311—Customer Relations Management Technical Assistance Services and served as the project director for the ICMA National Study of 311 and Customer Service Technology from 2006 to 2011. Fleming has written in various capacities about the use of data for improved local government service delivery and performance measurement. She currently manages the #LocalGov Technology Alliance, an Esri-ICMA initiative to explore the world of big data, open data, apps and dashboards, and what it all means for local governments. She served as editor of *The GIS Guide for Local Government Officials (2005)* and *The GIS Guide for Elected Officials (2014),* both joint publications produced by Esri and ICMA. Fleming has an extensive background in community development, having worked with local governments, community groups, and nonprofit organizations for more than a decade at the Iowa Department of Economic Development. Before joining ICMA, Fleming was with the Urban Institute's Center on Nonprofits and Philanthropy, where she conducted research on building the capacity of nonprofit organizations. Fleming received her BS from Iowa State University in community and regional planning and her MS from The Johns Hopkins University in environmental sciences.

**George Galster** is the Clarence Hilberry Professor of Urban Affairs at Wayne State University. He has held positions at the Universities of Harvard, Berkeley, North Carolina, Amsterdam, Delft, Glasgow, Mannheim, Oslo, and Western Sydney, among others. Galster served as director of housing research at the Urban Institute before coming to Wayne State University in 1996. His research has focused on urban neighborhoods and housing markets, exploring how they change and how they change the people who live within them. His latest book is *Driving Detroit: The Quest for Respect in the Motor City* (2012). He has provided housing policy consultations to public officials in Australia, Canada, China, Scotland, and the United States. He earned his PhD in economics from the Massachusetts Institute of Technology.

**Lina Hedman** is a researcher at the Institute for Housing and Urban Research, Uppsala University, and at the Institute for Future Studies, Stockholm, Sweden. Her main research interests are residential segregation, processes of neighborhood selection, and neighborhood effects. Most of her empirical research is carried out using Swedish register data, but she has also published papers discussing these issues in more general terms. She holds a PhD in human geography from Uppsala University.

**Julia Koschinsky** is an associate research professor at Arizona State University's School of Geographical Sciences and Urban Planning (SGSUP), where she is currently engaged in a national collaborative research project on the neighborhood context of subsidized housing. She is also the research director of the SGSUP GeoDa Center for Geospatial Analysis and Computation, which focuses on the development of spatial analytical methods and their implementation in free and open-source software programs. Koschinsky's research addresses questions at the intersection of spatial data analysis and evaluation methods in application areas such as housing, crime, and health. Her research and evaluation experience involving neighborhood-level data in the academic, business, and nonprofit sectors spans almost 20 years.

**Sheila Martin** is director of the Institute of Portland Metropolitan Studies and the Population Research Center and a faculty member in the Nohad A. Toulan School of Urban Studies and Planning at Portland State University. Prior to joining the university in 2004, Martin served as Washington State Governor Gary Locke's economic development advisor and as a senior economist at the Research Triangle Institute. Her research and teaching interests include sustainable economic development, regionalism, and technology economics and policy. Martin earned a BA in economics and political science from Southern Illinois University, an MA in international studies from the University of Kentucky, and a PhD in economics from Iowa State University.

**Jeff Matson** is director of Community GIS at the University of Minnesota's Center for Urban and Regional Affairs and a faculty member in the Master's of Geographic Information Science program. He has more than a dozen years of experience providing training, mapping, data, and technical assistance to community groups and nonprofits throughout the Twin Cities. Matson received a bachelor's degree in anthropology from the University of California, Berkeley, and a master's degree in geographic information science from the University of Minnesota, where his research focused on environmental justice and public participation GIS.

**Meg Merrick** is a research associate at the Institute of Portland Metropolitan Studies and assistant director for community and neighborhood geography in the College of Urban and Public Affairs at Portland State University. At the institute, Merrick coordinates Greater Portland Pulse and the Community Geography Project. She also serves as editor-in-chief of *Metroscape* magazine, the institute's flagship publication. Her research interests include exploring the ways in which maps and mapping empower community action and policymaking. Merrick holds an MS in geography and a PhD in urban studies from Portland State University.

**Lisa Nelson** is a senior policy analyst in the Community Development Office of the Federal Reserve Bank of Cleveland. Prior to joining the Federal Reserve in 2006, she worked in the Center on Urban Poverty and Community Development at Case Western Reserve University, where she contributed significantly to the development of the NEO CANDO data system. Nelson also worked with communities to use these data in their initiatives. Her recent work focuses on housing-related research, including analyses of loan performance, effects of local housing ordinances, and the effects of neighborhood stabilization efforts. She received a BA in political science from the University of Tennessee and an MPA from Cleveland State University.

**Greg Sanders** is a senior technical manager for Chapin Hall at the University of Chicago. He served as information architect for the Chicago Metropolitan Agency for Planning (CMAP) from 2002 to 2013. He was technical lead for CMAP's online data systems, including the MetroPulse applications. He cofounded the Illinois Data Exchange Affiliates data-sharing coalition in 2006 and the Full Circle community mapping and planning project in 2004. Sanders was previously information technology director for the Cleveland Housing Network and senior programmer–analyst for Penske Logistics. He has managed federal Technology Opportunities Program grants in both Cleveland and Chicago. Sanders holds a BA in political science from Oberlin College and a master's degree in mass communications from Kent State University.

**Jacob Wascalus** is a writer and analyst for the Community Development Department of the Federal Reserve Bank of Minneapolis. He writes about topics ranging from workforce development and neighborhood stabilization to suburban poverty and the housing market. He also works with colleges and neighborhood groups in designing their own data-oriented community development projects. Wascalus holds a master's degree in urban and regional planning from the University of Minnesota's Hubert H. Humphrey School of Public Affairs and a bachelor's degree in English literature from James Madison University.

# Index