Session 1: Wednesday 5/6/2015, 2pm-3pm

Location: Pittsburgh/Erie

Session Title: Using NCDB data

Organizer: Rob Pitingolo

Primary Notetaker: Rob Pitingolo

Participants: Rob Pitingolo, Mike Carnthan, Scott Gaul, Jennifer Bert, Jay Colbert, Sharon Kandris, Sean Capperis, Tyler Waring, Jane Downing, Richard Parks, Kim Pierson, Jeremy Pyne, Courtney Denton, Liz Monk, Whitney Soenksen, Young, Groesbeck, Ashley Clark, Mallory O’Brien, Susan Millea, Downing

Notes:

Carnathan: We did research on income inequality. Atlanta is leading the nation in inequality. We used it to look at those changes over time. Rather than talking about the city as a whole we are looking at the neighborhood level. There’s a lot more to the story than the national headline says.

Merrick: We have it but haven’t cracked it open yet. What is NCBD exactly?

Pitingolo: NCBD allows longitudinal analysis of Census ACS data. Since Census tracts change every decade, you can’t compare one decade to another, NCBD fixes that.

Capperis: We have done a few things. First thing is ripping it out of the software. We look at NYC, the Metro tracts and then the US as a whole. Our data analyst has come up with a Stata process that takes it from the native wide format and turns it into a long format which is more useful for analysis. In theory should transfer to SAS with good fidelity. This was a big undertaking and we’re happy to share. On the research side we looked at patterns of density since the 70s. Main measure is not raw density but experience density - the median of the density weighted by population in the tract. We charted that over time because a big thing happening in NY is the idea that neighborhoods can’t handle growth, so there needs to be rezoning to grow. We are trying to add facts to this raging debate. Another project that we’re starting up is a new research agenda looking at gentrification overall and how that affects holders of housing choice vouchers at the neighborhood level and building level.

Colbert: Professor at our university brought up Chicago gentrification index. Look at Chicago neighborhoods and see how neighborhoods change over time to see if they are getting better or experiencing gentrification. Looked like you can get most of the variables out of NCDB.

[unknown]: I want to look at Greenfield development. I can see how allocating data to certain census tracts.

Carnathan: Untracted areas aren’t included in 70 and 80, some process was used to correct. There are flags for how it has changed over time.

Pyne: What about how changes in questions.

Capparis: There are some fields where they make the connection for you, but you have to do the work on the universe yourself. They will report straight through but you have to watch out and understand the differences. There are places where you won’t come up with same results from the summary tables. Summing up by total population won’t always add up but you can see where that happens.

Colbert: is there any geography other than tracts?

Capperis: Just tracts. It should nest properly but not sure. I do have a question about CBSAs. Not sure if you can compare them over time.

Carnathan: Not all areas of ATL tracted in 1980.

Barndt: Why was this developed by UI? Is it because of work on poverty?

Pitingolo: NCDB has been used heavily in work on concentrated poverty. It has been used outside of UI. You can find many reports and phd dissertations written if you search for NCBD as a data source.

Gaul: I tried to use iPUMS and failed. Can I track a single person longitudinally. I want to know if Mrs. Jones moved from a high poverty neighborhood to somewhere else.

Capparis: Can go to a census data center to get at some of that information. Need to get special status.

Pitingolo: Right, NCBD can’t do that.

Kandris: With Census, there is so much there. You can take any variable and do any analysis. What are topics that are interesting to look at.

Carnathan: We used a student to go through and document what we have available. Educational attainment is an easy one to get started with.

Barnt: The age adjusted data gets quite interesting if you do detailed histograms over time. What you get when you do that is the sense of what kind of families move in and out. Where do young women and children show up. Gets even more interesting when you can get into block level data. Seeing entire groups of people coming and going in a neighborhood can be easily visualized.

Young: Can anyone get this?

Pitingolo: It’s for sale by Geolytics. NNIP partners get a complimentary copy.

Carnathan: There’s also a brown longitudinal database.

Barnt: Is the process from Brown and NCBD different?

Pitingolo: From what I can tell the process is similar. Brown has a limited number of variables. NCBD has every variable. I am working on a more detailed writeup on this.

Capparis: Would love to have a process to pull in all the new ACS data as it is released. We want something more contemporary to short form data.

Kandris: Are there restrictions for republishing.

Colbert: There are limits. There are restrictions.

Carnathan: What would be the restrictions in just plugging in new data? Is that bad?

Capperis: I have a strong opinion of this based on my personal usage. When possible I try to avoid using the ACS. I would probably do it with a grain of salt and use the short form when it’s there. There’s no reason you couldn’t just stick on the 5 year and run with it.

Carnathan: That’s what we’re doing.

Capperis: We use NCBD to look at counts and that’s where I start to get scared.

Pyne: As new ACS come out, I would keep using the newest 5 years.

Barnt: So basically the confidence interval situation is still there, right?

Pitingolo: Any problem with the ACS is a problem with NCDB.

Carnathan: Depends on the variable, some better than others. The sample was still better in 2000 but still had issues.

Young: For SLC we did do our own 2000 to 2010 change. We do have numbers that are just SLC boundary. Can do that to match a place boundary. Untracted areas are something that you need to be cautious about.

Kandris: We have a city-wide festival done each year. They adopt a theme. Organizations have events around that theme. Next year’s theme is home. The idea is something that adopts a neighborhood or two and how has that changed over time. We want to do a better job of telling stories with our data and we can do that.

Capperis: Housing demolition is one component in our piece. There are entire neighborhoods that have hollowed out. Neighborhoods went from high to low density.

Downing: I used previous Census to track racial change in Pittsburgh. I haven’t done it recently. We have minor immigration here and we need to look at it again. Public policy actually changed the composition of many neighborhoods. It tells a different story than the aggregate that we are telling now.

Kandris: Our newspaper did that a few years ago. Similar to the inequality index. Saw shifting race over time. Could be expanded over time.

Bardt: Milwaukee has similar experiences and Cleveland and Detroit. The relationship between housing patterns and population patterns. Property data isn’t good at understanding vacancies. If you’re dealing historically, people have forgotten what it’s like in the 60s - very high densities in a property. A neighborhood might lose population without losing housing or increasing vacancy. How can a neighborhood grow. A good housing dataset works rather nicely with the census.

Capperis: Is there infrastructure to continue the conversation about this to keep it going?

Kandris: How can we share the data in a way that’s collaborative?

Capperis: We are trying to setup a github at the Furman Center. We can try to blast out something to partners. Along with other processes.

Kandris: This isn’t the cross-site session, but it would be great to do a combined report.

Millea: Suburbanization of poverty. How to deal with it.

Carnathan: In Atlanta, 8 counties were tracted in 1970 out of 20+ today. NCBD isn’t really good for work in suburbs.

Pitingolo: Anyone else? OK we’ll break.

[end of session]