

***The Effects of Foreclosure on
Children and Schools:
A NNIP Cross Site Study***

***New York City Analysis Plan
and Data Diagnostic Memo***

**Prepared by the Furman Center for Real Estate and Urban Policy
New York University School of Law
for The Urban Institute**

January 2010

This is a planning document prepared for a project funded by the Foundation for Open Society that will examine the effects of foreclosures on public school children in three cities (Baltimore, New York, and Washington, D.C.). The working products are being made available as a resource for organizations who are interested in doing similar analysis for other cities.

Please do not cite or quote this document without permission of the authors.

For more information about the overall project, see <http://www2.urban.org/nnip/foreclosures.html> or contact the project coordinators: Kathy Pettit (kpettit@urban.org) and Jenn Comey (jcomey@urban.org)

Foreclosure and Student Mobility Context

a. *Foreclosure Trends Context.*

New York City's housing market has unique scale, density, and diversity. Over 8 million New Yorkers live in a mixture of condominium and cooperative apartments, 2-to-4 family buildings, and single family homes across the five boroughs, comprising over 3.3 million housing units. The homeownership rate in New York City is 33.6%, significantly lower than the national rate (67.2%) and rates for comparable cities (Boston 38.5%, Chicago 49.9%, Philadelphia 57.5%, Washington, D.C. 44.5%).¹ The rental vacancy rate in the city has hovered around 3.5% since 2000, and renter households face a median rent burden of almost 30%.² More than half of the rental units in New York City are rent-regulated and 14.8% of rental units are subsidized.³

Since 1974, New York City's housing market has experienced two large booms and two relatively small busts, resulting in a housing price increase of 250% from 1974 to 2006.⁴ Over the past few years, however, effects of the national economic downturn have become evident in New York City. The number of residential units authorized by new building permits in New York City declined in 2006 for the first time in ten years, foreclosure filings increased by 50% between 2006 and 2007, prices of single family and 2-to-4 family properties fell by over 13%, and rates of home purchase mortgage and refinance mortgage originations decreased dramatically.

The number of properties receiving a notice of foreclosure (*lis pendens*) increased by over 100% in New York City between 2000 and 2008, with the sharpest upturn occurring between 2005 and 2007 (see Table 1, Figure 1). In 2008, almost 32,000 residential units were located in properties that received a notice of foreclosure.

The impact of this current wave of foreclosure activity is not borne solely by homeowners. Of the nearly 15,000 properties that received a *lis pendens* in 2008, over half were multifamily properties (see Figure 2). While the number of foreclosures of single family properties has grown considerably since 2000 (82%), the number of *lis pendens* for 2-4 family properties has increased at a much greater rate (144%). Rental units make up approximately 53% of the units in properties that received a *lis pendens* in 2008 (see Table 3).⁵ Tenants living in properties that enter foreclosure have few protections; and if a property is sold at auction, most tenants face eviction.

¹ American Community Survey 2007.

² American Community Survey 2007.

³ NYC Housing and Vacancy Survey 2005.

⁴ Armstrong, A., et al., State of New York City's Housing and Neighborhoods Report: 2008

⁵ This analysis (Furman Center) makes a number of conservative assumptions. We assume that all single family and condo units are owner-occupied. We assume that an owner lives in one of the units in all 2-4 family buildings (and that renters occupy the remaining units). And we assume that all units in 5+ apartment buildings are occupied by renters.

Foreclosures in New York City are geographically concentrated in the outer boroughs (see Figure 3). Brooklyn and Queens have experienced dramatic increases in the number of properties receiving *lis pendens* filings since 2000, while the Bronx, Staten Island, and Manhattan have shown slight, but steady increases. The bulk of renters affected by foreclosures are also concentrated in Brooklyn and Queens.

After a *lis pendens* filing, 8% of properties are sold in arms-length sales within 3 months, 27% are sold within one year, and almost half are sold after 5 years (see Figure 4). One year after the *lis pendens* filing, only 2% of properties have been sold at auction, 5% were transferred for other reasons (such as a divorce settlement, non-arms-length sale, etc.), and over 3% have received a subsequent *lis pendens* filing. Five years after the *lis pendens* filing, almost 8% of properties have been sold at auction, 11% were other transfers, and 11% have received subsequent *lis pendens* filings.

**Table 1. Properties that
Received a Lis Pendens in
New York City**

Year	Number of Properties
2000	7,350
2001	7,297
2002	7,977
2003	7,623
2004	6,940
2005	6,869
2006	9,704
2007	14,525
2008	14,798

Figure 1. Properties that Received a Lis Pendens Filing (2000-2008)

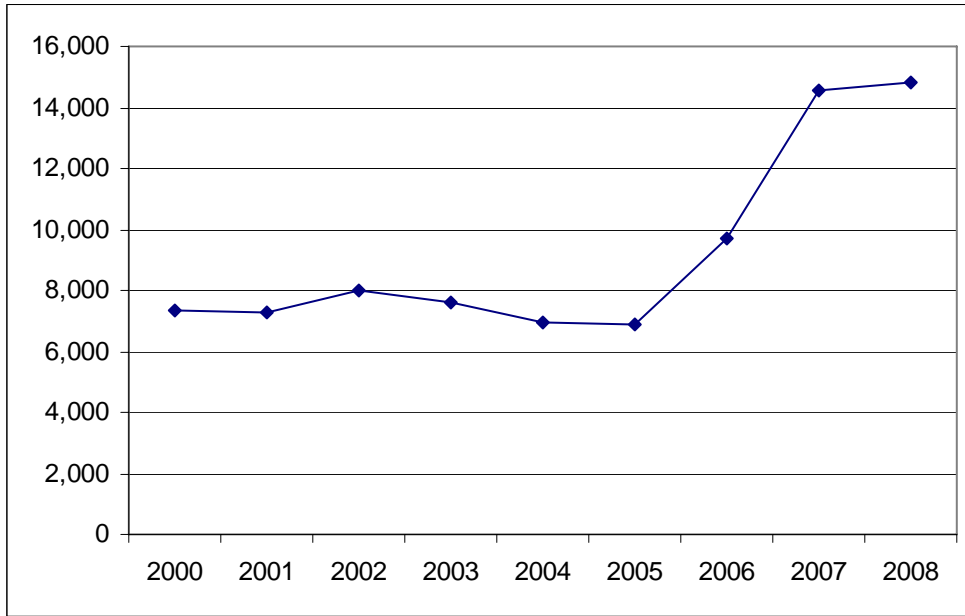


Figure 2. Properties that Received a Lis Pendens Filing, by Housing Type (2008)

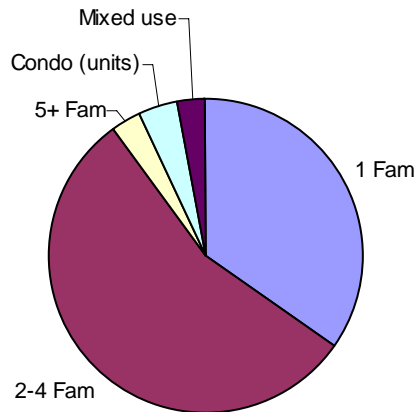


Table 2. Properties that Received a Lis Pendens by Housing Type

Year	1 Fam	2-4 Fam	5+ Fam	Condo (units)	Mixed use
2000	2,810	3,353	507	327	353
2001	2,876	3,382	408	299	332
2002	3,102	3,809	342	390	334
2003	2,970	3,701	348	296	308
2004	2,767	3,302	256	353	262
2005	2,685	3,401	235	323	225
2006	3,605	5,118	251	432	298
2007	5,103	8,128	345	591	358
2008	5,128	8,177	454	625	414

Table 3. Estimated Rental Units in Properties that Received a Lis Pendens

Year	Estimated Number of Rental Units in LP Properties	% of Total Units in LP properties that are Rental Units
2000	9,372	52.8%
2001	10,988	59.0%
2002	8,496	51.2%
2003	8,999	53.6%
2004	6,770	48.0%
2005	6,835	48.8%
2006	9,205	47.8%
2007	14,643	49.9%
2008	16,639	52.6%

Figure 3. Notices of Foreclosure, NYC, 1-4 Family Properties (2008)

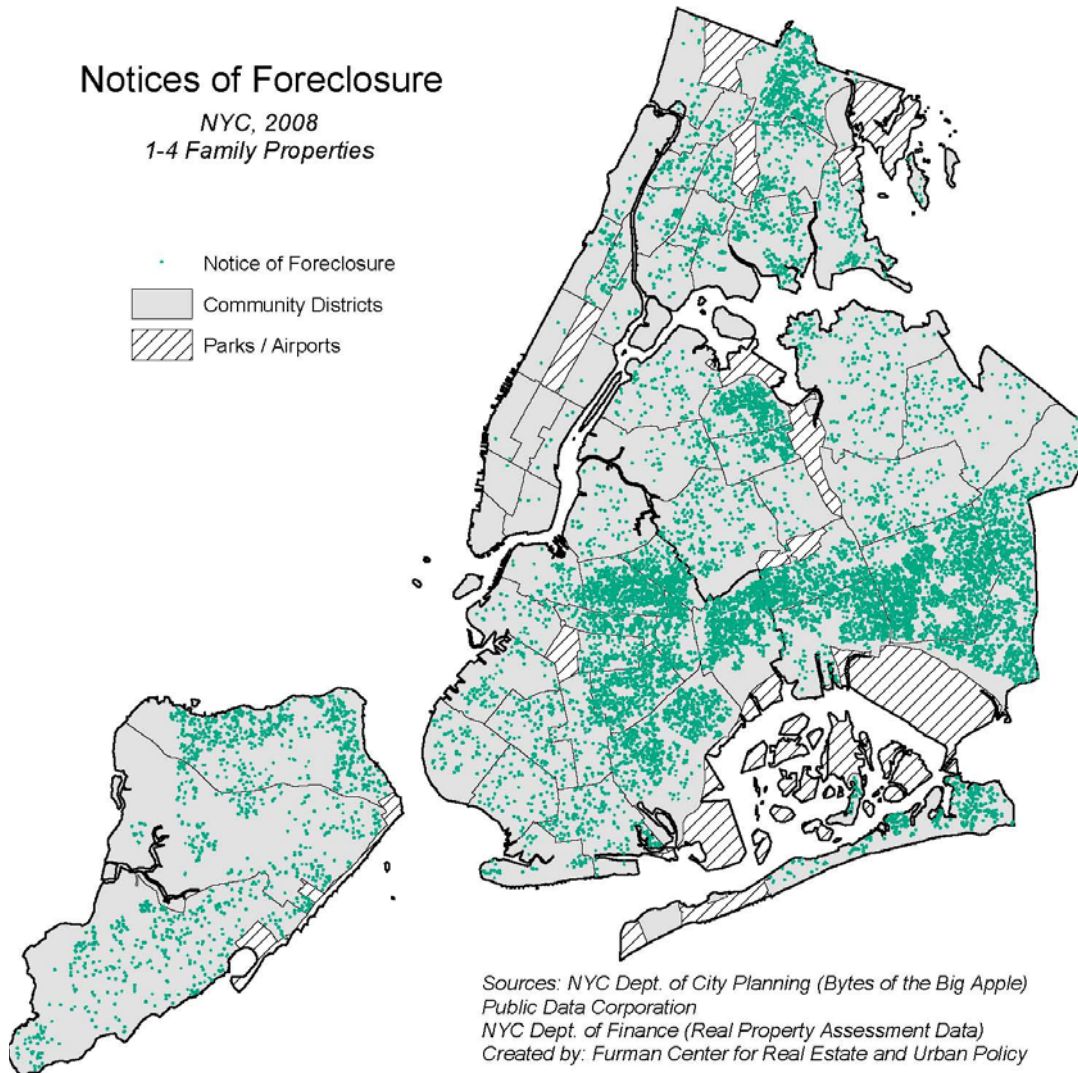
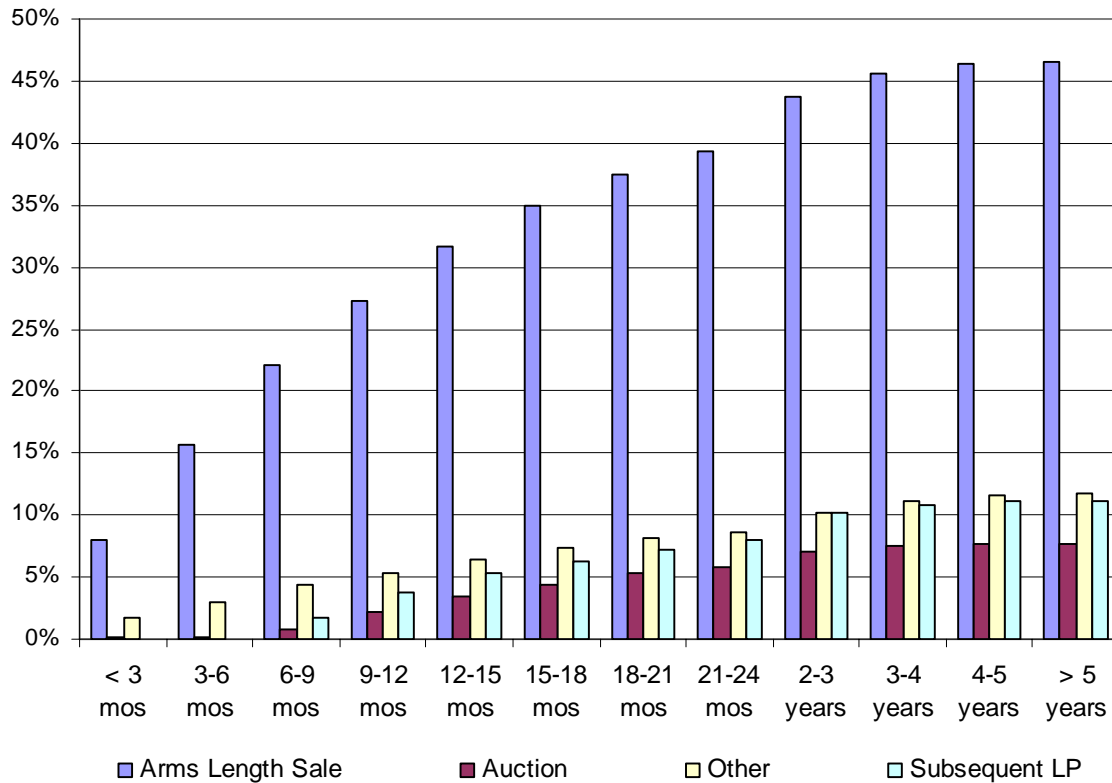


Figure 4. Cumulative Outcomes of 2003 *Lis Pendens* Filings⁶



b. Student Mobility Context.

Previous research has found that students in New York City primary schools experience considerable mobility.⁷ Approximately 12% of New York City primary school students switch schools at times other than the major endings of primary school in a given year. Black and Hispanic students tend to switch schools between school years at higher rates than white or Asian students before the 5th grade, and poor students switch at higher rates than non-poor students. Although moves to a different school during the school year are more rare (only 6.7% of students), black and Hispanic students are also more likely to experience this type of school change than white and Asian students. Looking at the experiences of a cohort of 8th grade students, the study found that a quarter of the students who have been continuously registered since 5th grade attended three schools by the 8th grade. Almost half of the students in the 3rd grade cohort who moved relocated to a new zip code, and 17% moved to a different borough, while students who moved in the 6th

⁶ Values do not sum to 100% because properties with no outcome yet are not included in the chart.

⁷ Amy Ellen Schwartz, Leanna Stiefel, and Luis Chalico. (2007) “The Multiple Dimensions of Student Mobility and Implications for Academic Performance: Evidence from New York City Elementary and Middle School Students.” A Condition Report for the New York Education Finance Research Consortium.

grade cohort were less likely to move to a new zip code. More student mobility, measured by the number of schools attended, has been found to have a negative effect on 8th grade reading scores.⁸

Phase I: Students experiencing foreclosures in 2004 and 2008

Research Questions

- How many public school children have been affected by foreclosures in New York City? Has the number changed over time?
- What are the social and demographic characteristics of the school children who are affected by foreclosure (e.g. race/ethnicity, gender, grade, free/reduced price lunch, LEP, special education, country of birth)? Have they changed over time?
- Are these foreclosed students clustered in particular schools/neighborhoods? Trends in school and neighborhood characteristics 2004 vs. 2008.

Analytic Work

Data sets

In order to answer the Phase I research questions, we will use the following data sets about public school students, schools and foreclosures.

Public school student data

- Source: The student-level data are provided by the New York City Department of Education (NYCDOE).
- Annual student data are provided for all students enrolled on October 31 of each year. Addresses are also provided for all students enrolled on March 1 and June 1 of each year.
- The student-level data include the following variables: grade, date of admission, country of birth, race/ethnicity, gender, free and reduced price lunch status, Home Language, Limited English Language (LEP), special education status, school attending.
- Addresses for properties that received a *lis pendens* filing were compiled by NYU and given to the NYC DOE to match to student address records. See “Merged Student and Foreclosure dataset” below.

⁸ The coefficient on another measure of mobility – number of moving years – was negative but initially insignificant. With the inclusion of school fixed effects, the coefficients maintained the negative signs, but the mobility measure of the number of schools attended lost its significance, while the number of moving years gained significance and magnitude.

Public school data

- Source: NYC DOE Annual School Reports (ASRS) and School Based Budget Reports (SBERS) as cleaned by Institute for Education and Social Policy
- The school level data include variables on demographics of the school (e.g. share of students by race/ethnicity, free and reduced price lunch, LEP, percent immigrants, gender), resources (e.g. expenditures, teacher characteristics), and average test scores and attendance data.

Foreclosure data

- Source: Lis Pendens filings from the Public Data Corporation, updated quarterly (1997-2008); Sales data from the New York City Department of Finance, updated annually (1997-2008).
- Unduplicated foreclosure notices for properties by structure type (limited to residential properties)
 - Residential, single-family home
 - Residential, condo
 - Residential, cooperative
 - Residential, rental apartment building
- The foreclosure data include foreclosures that originated in 2002 up until 2008.
- By matching the foreclosure notice data to property sales records and mortgage lending data, the Furman Center has created a file with records for each foreclosure “episode” (property/notice period) and assigned an outcome for the property.
 - 1-In Foreclosure (no outcome following LP)
 - 2-Auction
 - 3-REO
 - 4-Arms-Length Sale
 - 5-Other Transfer
 - 6-Subsequent LP
- The properties with foreclosure outcomes should have a corresponding first notice date (date of which the first foreclosure notice was recorded) and outcome date (or the date of which the outcome was recorded).
 - Properties in the foreclosure process would not have an outcome date yet.

Merged student and foreclosure data

- 1) Matched foreclosure data to student data by parcel (or at property level).

To match students to properties receiving foreclosure notices, the NYCDOE will match the students' addresses, for the 2003-2004 and 2006-2007 school years, to the LP addresses. For each of those school years, we will link student data to their records from the previous school year and the following school year(s). For instance, we will link students in 2003-2004 to their records in the previous year (2002-2003) and the following school year (2004-2005) to allow for a pre-post comparison.

As is inevitable in empirical research, we have to use some approximations in identifying students whose families go through a foreclosure in a given academic year (which we define as starting during the previous summer). The companion data diagnostic memo goes into more detail on the choices we have made to balance the potential sources of measurement error. We rely on two alternative definitions. Our more conservative alternative flags the following three groups of students as experiencing a foreclosure during the school year 2003-2004:

- a. Students who live in a property on June 1st, 2003, which receives an LP between June 1st, 2003 and Oct 31st, 2003;
- b. Students who live in property on Oct 31st, 2003, which receives an LP between Oct 31st, 2003 and March 1st, 2004; and
- c. Students who live in a property on March 1st, 2004, which receives an LP between March 1st, 2004 and June 1st, 2004.

This definition is conservative for several reasons. First, because we only count foreclosure notices that are issued after the date on which we know a child's address, we can be certain that the foreclosure was not resolved or completed before the student's family moved into the building. Second, because we only match students to LPs issued in the four months after we capture a student's address, we can be fairly confident that the student's family was still living in the property at the time the notice was issued.

To be clear, we are not assuming that foreclosures only last 3-5 months. In New York City, the foreclosure process is slow and takes about 18 months on average. The windows we are defining are simply about matching students to foreclosures. In our analysis, we will test for longer-term impacts of foreclosure. So for instance, we will test whether a student's mobility across schools during the 2003 and 2004 school years is affected by her living in a property going through foreclosure during any of the three

windows above (between June and October 2003, November 2003 through Feb 2004, and March through May 2004).

This definition may be overly conservative and miss students who move into a unit after a foreclosure is filed, but before a foreclosure is resolved. So, for example, if a property receives an LP in July, and a student moves into that property in August, we will not count them as affected. Yet the foreclosure may not yet be resolved. Our second definition attempts to include these children.

Through this matching process, we will create a student-level data set with corresponding information about whether that student lived in a property that went through foreclosure during that academic year or the previous academic year (using our two definitions and indicating the window when foreclosure started). We will also include in the file the foreclosure outcome. Not all *lis pendens* result in a foreclosure auction, because the borrower may be able to resolve the delinquency or modify the mortgage in order to keep the home. Accordingly, we will look separately at children living in properties that sell at auctions, go to REO, or sell through arms-length sales within 12 months of the foreclosure notice.

Neighborhood data. Census tract level data is available from the GeoLytics Neighborhood Change Database for the census years 1990 and 2000, and Public Use Microdata Area (PUMA) level data is available for more recent years from the American Community Survey. Data compiled by the Furman Center will also be used, including racial and income diversity indices and a repeat sales index.

- Source: GeoLytics Neighborhood Change Database (1990, 2000), American Community Survey (2005-2007), Furman Center
- To proxy for neighborhoods, we use elementary school zones. Census tract data from the Neighborhood Change database has been weighted and aggregated to the school zone level. Available variables include poverty status, immigrant status, racial composition, family composition, unemployment, and educational attainment, among others.

Analysis Plan

- We will first analyze the basic descriptions of the public school students affected by foreclosure.
 - Number/share of public school students affected by foreclosures for 2003-04 and 2006-07
 - Characteristics of affected students for 2003-2004 and 2006-2007

- E.g. Race, gender, foreign-born, grade, free/reduced price lunch, special education, LEP.
- Any difference across foreclosure outcomes? Especially renter vs homeowner?
- School analysis of affected students
 - If there is a concentration in specific schools, what are the characteristics of the schools? (Location, average test score, other)
 - Identify differences between the two school years
- Neighborhood analysis of affected students
 - Are students concentrated in particular neighborhoods? Characteristics of neighborhoods.
 - Identify differences between the two school years

Phase II: Students experiencing foreclosures versus ones not experiencing foreclosure, pre and post

Research Questions

- How do students who go through foreclosure compare to those who do not both pre and post foreclosure?
- Do students affected by foreclosure switch schools post foreclosure outcome? Compare to non-foreclosed students.
- What schools do foreclosed and non-foreclosed students enroll in post foreclosure year? What are their characteristics? How do they differ from the previous school?
- How do neighborhoods for elementary school students foreclosed and not compare pre and post foreclosure years?

Analytic Work

Data sets

In order to answer the Phase II research questions, we will use the following data sets about public school students, foreclosure data, and neighborhood and school characteristics.

Public school student data. We will be using the same student-level described in Phase I. We will also use data from the previous year for all students to compare baseline, pre-foreclosure test scores and mobility patterns among kids who go through foreclosure, and

data from the following year to compare outcomes after foreclosure (i.e. for the 2003-04 school year, we will look at outcomes in 2002-03 and 2004-05 and for the 2006-07 school year, we will consider student outcomes in 2005-06 and 2007-08).

Public school data. We are using the same school level data as described in Phase I.

Foreclosure data. We are using the same foreclosure data as described in Phase I.

Neighborhood Data. We are using the same neighborhood data as described in Phase I.

Analysis Plan

Relying on the matched student and foreclosure data for 2003-2004 and 2006-2007, plus linked student data for the pre and post years – 2002-03 and 2004-05; 2005-06 and 2007-08:

- We will describe baseline, pre-foreclosure and post-foreclosure test scores and other characteristics of students who go through foreclosure and those that do not (comparison group).
- We will analyze foreclosed and non-foreclosed students (same years, comparison group) to see if mobility across schools differs between foreclosed year and post year and between foreclosed and non-foreclosed students.
 - For those foreclosed and non foreclosed students who switched schools, we will compare the characteristics of the schools pre and post move.
 - We will break the analysis down into those foreclosed students who moved and who didn't move.
 - We will analyze demographics, foreclosure outcome homeowner vs. renter, and not foreclosed, and neighborhood characteristic of students who moved school.
 - Challenges:
Students who reach the final grade in the school are required to switch schools. Therefore, we will identify the students who have reached their maximum grade at their school versus those who have not. The analysis will separate these two groups.

- We will compare schools and neighborhoods of foreclosed students and non foreclosed students in pre and post foreclosure years.
- Challenges for all of phase II:
 - This analysis will only capture those students that remain in the public school system and will not include students who have left altogether.
 - Some students in the comparison groups in for the pre and post-years will be foreclosed in those years and we may not know that.

Stakeholders

Analysis plan review

We consulted with a number of experts in the field in designing our analysis plan. We spoke to Tom Gold, Director of External Research, Evaluation and Reporting at the New York City Department of Education about the student-level data, the address information available, and the feasibility of the match. We spoke to Sarah Gerecke while she was CEO of NYC Neighborhood Housing Services about the timing of the foreclosure process and likely impacts on families and children. We also spoke to Ellen Howard-Cooper, Deputy Commissioner for Policy and Planning at the New York City Department of Homeless Services, to learn more about the families who were coming to the shelter system from foreclosed buildings. Finally, we discussed our analysis plan with a few other researchers including Peter Messeri and Brendan O’Flaherty at Columbia University.

Stakeholder review

We plan to ask Elyzabeth Gaumer, who is Director of Research at the New York City Department of Housing Preservation and Development to review our findings. We will also share drafts with key practitioners who work in the field doing foreclosure prevention and assistance. We have already consulted with Jessica Attie, who is co-Director of the Foreclosure Prevention Project at South Brooklyn Legal Services about the foreclosure process and plan to review our findings with her as well as with her colleague Meaghan Faux. As for education stakeholders, we will share our results with Tom Gold from the Department of Education, as well as with Jim Liebman, who is a Professor of Law at Columbia University and until recently was the Chief Accountability Officer for the City’s Department of Education.

Data Approval. New York University has received approval from NYCDOE to use their student-level data for the OSI study.

School Data

Student-level data from the New York City Department of Education (NYCDOE) are updated annually for all students enrolled on October 31 of each year. Demographic variables included in the student-level data are grade, date and country of birth, race, ethnicity, gender, free and reduced price lunch status, and home language. The data set also includes information on annual school attendance, Limited English Proficient (LEP) status, special education status, and standardized test scores, which are taken in the winter and spring of each academic year. Finally, we have each student's address at three times during the school year: October 31st, March 1st and June 1st. Data for individual students can be linked from one academic year to another, as long as the students remain within the New York City public schools.

Data on individual students can also be linked to school-level files, which include the demographics of the school (such as the share of students by race, ethnicity, and gender, the share qualified to receive free and reduced price lunch, the percentage who are LEP, and the percentage who are foreign-born), as well as resource data (expenditures, teacher characteristics) and average test score and attendance data.

For each year, we exclude students who are missing admit/discharge dates, school, or grade codes. Excluding students with missing grade codes decreases the number of full-time special education students in our dataset because these students typically are not assigned to a graded classroom. In most cases these students would be dropped from the analysis anyway, because they do not typically have test score data. Also, the majority of these students attend schools in District 75, which is designed specifically for students with special needs, and there is limited information available about these schools.

Because DOE occasionally makes changes in the data provided, we have developed a standardized protocol to enable year-to-year comparisons. For example, the variables provided to designate a student as special education or LEP have changed over time because of changes in reporting requirements. Based on the information provided we create a flag that designates a student as special education or LEP.

Foreclosure Data

We use a dataset of parcel level *lis pendens* (LP) filings from the Public Data Corporation. This dataset is updated quarterly and available from 1997 to 2008. The dataset includes all residential parcels that received a notice of foreclosure and the date of each notice. The final file used for the match to the NYCDOE student data includes the date and borough-block-lot (BBL) identifier for each filing, as well as street address, property characteristics, and information about the disposition of the property after the LP was issued (e.g., whether the property was sold at auction, whether it was transferred in an arms-length sale, whether it received a subsequent *lis pendens* filing).

To match the foreclosure data to the student-level data, we identified the street address of each foreclosure filing.¹ For properties that had more than 1 LP between sales, and the LP is

¹ Details about the matching process: First, we created an LP file with 1 observation per BBL per year, for LPs occurring between 2002 and 2008 (65,806 observations). Next, we used PLUTO to create a file that of

within 6 months of the prior filing, we consider it to be the same "instance of foreclosure" and trace outcomes based on the date of the first LP. LPs that occur more than six months after an earlier LP are included and considered a separate foreclosure filing. In total, the final dataset includes 65,793 *lis pendens* filings with addresses. Condominium units are considered to have one residential unit. Definitions of the variables in the final dataset are presented in Table 1, and Table 2 presents descriptive statistics.

Each property with a notice of foreclosure is tracked over time and assigned a foreclosure outcome. The outcomes range from the sale of the property at auction to properties that have not been sold one year later. Table 3 below lists the outcomes of the LPs in our sample within one year of the filing date, and the share of properties with each outcome since 2002. Since many foreclosures are not resolved in the first year following a *lis pendens* filing, the outcomes of LPs within three years of the filing date are shown in Table 4.

Table 1. Final Variable List & Definitions

Variable Name	Definition
Borough	Borough Name
ADDRESS	Address (Unit Number and Street Name)
add_num	Address Number alone
add_strt	Street Name alone
ZIPCODE	Zip Code
Tract00	Census Tract (use with borough code)
RES_UNIT	Number of Units
LPDATE_1	Date of <i>lis pendens</i> filing
LPoutcome	Code identifying potential outcomes (e.g., sold at auction, arms-length sale)
ddate_p	Date of Sale Prior to <i>lis pendens</i> filing
ddate_s1	Date of First Outcome After <i>lis pendens</i> filing
Time2nextLP	Number of days until subsequent <i>lis pendens</i> filing

Table 2. Descriptive Statistics (lp_0208_FINAL.sas7bdat)

LPs between 2002-2008					
Variable	N	Missing	Min	Max	Mean
ZIPCODE	65,793	0	10,001	11,694	11,083
tract00	62,473	3,320	1	1,621	481
RES_UNIT	65,793	0	1	1,255	2
lpdate_1	65,793	0	15,342	17,897	16,804
ddate_p	63,418	2,375	5,204	17,864	14,921
ddate_s1	29,510	36,283	15,343	17,897	16,809
time2nextlp	9,574	56,219	184	2,576	627

street addresses, with one observation per BBL. We merge this file with the LP data set by BBL (62,633 observations, 95% match). Observations that did not match to an address in PLUTO (3,173 observations) were matched to RPAD address data (3,134 observations, 99% match). Next, we removed the 63 remaining observations that did not match to an address and the 157 observations with building classes G0 and Z0 (which indicate nonresidential uses such as garages & tennis courts, etc.), resulting in a final dataset of 65,793 observations.

Table 3. Distribution of LP Outcomes within 12 Months of LP Filing

	2002		2003		2004		2005		2006		2007		2008		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1-Auction	89	1%	67	1%	75	1%	80	1%	92	1%	96	1%	16	0%	515	1%
2-REO	99	1%	83	1%	60	1%	93	1%	324	4%	644	5%	122	1%	1425	2%
4-Other Transfer	361	5%	394	5%	331	5%	294	4%	360	4%	324	2%	165	1%	2229	3%
5-Arms length sale	1583	20%	1893	26%	1999	30%	1910	29%	2280	25%	1471	11%	663	5%	11799	18%
6-Subsequent LP	426	5%	382	5%	347	5%	402	6%	561	6%	817	6%	235	2%	3170	5%
7-No Subsequent Outcome	5193	67%	4520	62%	3875	58%	3842	58%	5613	61%	10593	76%	13019	92%	46655	71%
Total	7751		7339		6687		6621		9230		13945		14220		65793	

Table 4. Distribution of LP Outcomes within 3 Years of LP Filing²

	2002		2003		2004		2005		2006		2007		2008		Total	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1-Auction	413	5%	312	4%	266	4%	230	3%	277	3%	184	1%	16	0%	1698	3%
2-REO	282	4%	180	2%	146	2%	289	4%	1011	11%	1140	8%	122	1%	3170	5%
4-Other Transfer	803	10%	730	10%	581	9%	520	8%	528	6%	396	3%	165	1%	3723	6%
5-Arms length sale	2865	37%	3036	41%	2791	42%	2521	38%	2829	31%	1807	13%	663	5%	16512	25%
6-Subsequent LP	1277	16%	1148	16%	1164	17%	1290	19%	1587	17%	1524	11%	293	2%	8283	13%
7-No Subsequent Outcome	2111	27%	1933	26%	1739	26%	1771	27%	2998	32%	8894	64%	12961	91%	32407	49%
Total	7751		7339		6687		6621		9230		13945		14220		65793	

Matching Students to Residential Parcels in Foreclosure

To match students to properties receiving foreclosure notices, the NYCDOE will match the students' addresses, for the 2003-2004 and 2006-2007 school years, to the LP addresses. For each of those school years, we will link student data to their records from the previous school year and the following school year(s). For instance, we will link students in 2003-2004 to their records in the previous year (2002-2003) and the following school year (2004-2005) to allow for a pre-post comparison.

As is inevitable in empirical research, we have to use some approximations in identifying students whose families go through a foreclosure in a given academic year (which we define as starting during the previous summer). For one thing, we only observe a student's address at three different times during each academic year, on June 1st, October 31st, and March 1st. Thus, if a student's family moves between June and October, we will not know exactly when they moved. At the same time, we are not always exactly sure when a foreclosure is resolved, which is one reason why we use the start of a foreclosure as our key measure of foreclosure.³ If, for

² The distribution of 3 year outcomes will surely change for *lis pendens* filed in 2007 and 2008 as we collect data in the coming years.

³ For more than half of our LPs, we see no further action over a 12-month period. Note however, that in our analysis, we will consider a foreclosure as resolved if a property is subsequently sold. We will also test whether students are differentially affected by foreclosures that end in different outcomes (e.g., property sold at auction; property ends up in REO; property sold in an arms-length sale; no further action).

instance, we see no subsequent sale or LP for a property within a year after the initial LP is issued, we cannot be sure whether the default has been resolved, the lender is forbearing on the foreclosure because the owner is still struggling to get current on his or her mortgage, or the foreclosure is just working its way through the process slowly. Depending on the windows we use to define foreclosure, these uncertainties can lead to under or over-counting of the number of kids going through foreclosure.

We balance these different sources of measurement error by using two different definitions of foreclosure – a conservative definition and a more liberal one:

The More Conservative Definition of Children Facing Foreclosure: For the conservative alternative, we flag the following three groups of students as experiencing a foreclosure during the school year 2003-2004:

- a. Students who live in a property on June 1st, 2003, which receives an LP between June 1st, 2003 and Oct 31st, 2003;
- b. Students who live in property on Oct 31st, 2003, which receives an LP between Oct 31st, 2003 and March 1st, 2004; and
- c. Students who live in a property on March 1st, 2004, which receives an LP between March 1st, 2004 and June 1st, 2004.

This definition is conservative for several reasons. First, because we only count foreclosure notices that are issued after the date on which we know a child's address, we can be certain that the foreclosure was not resolved or completed before the student's family moved into the building. Second, because we only match students to LPs issued in the four months after we capture a student's address, we can be fairly confident that the student's family was still living in the property at the time the notice was issued.

To be clear, we are not assuming that foreclosures only last 3-5 months. In New York City, the foreclosure process is slow and takes about 18 months on average. The windows we are defining are simply about matching students to foreclosures. In our analysis, we will test for longer-term impacts of foreclosure. So for instance, we will test whether a student's mobility across schools during the 2003 and 2004 school years is affected by her living in a property going through foreclosure during any of the three windows above (between June and October 2003, November 2003 through Feb 2004, and March through May 2004). Figure 1 summarizes the conservative approach.

This definition may be overly conservative and miss students who move into a unit after a foreclosure is filed, but before a foreclosure is resolved. So, for example, if a property receives an LP in July, and a student moves into that property in August, we will not count them as affected. Yet the foreclosure may not yet be resolved. Our second definition attempts to include these children.

More Liberal Definition of Children Facing Foreclosure: For the liberal alternative, we expand our definition slightly and flag the following three groups of students as experiencing a foreclosure during the school year 2003-2004:

- a. Students who live in a property on June 1st, 2003, which receives an LP between June 1st, 2003 and Oct 31st, 2003;⁴
- b. Students who live in property on Oct 31st, 2003, which receives an LP between June 1st, 2003 and March 1st, 2004, and which is not sold between the time the LP is issued and Oct 31st, 2003;
- c. Students who live in a property on March 1st, 2004, which receives an LP between Oct 31st, 2003 and June 1st, 2004, and which is not sold between the time the LP is issued and March 1st, 2004;
- d. Students who live in a property on June 1st, 2004, which receives an LP between March 1st, 2004 and June 1st, 2004, and which is not sold between the time the LP is issued and June 1st.

Figure 2 shows the students identified as going through foreclosure, using this more liberal definition.

Through this matching process, we will create a student-level data set with corresponding information about whether that student lived in a property that went through foreclosure during that academic year or the previous academic year (using our two definitions and indicating the window when foreclosure started). We will also include in the file the foreclosure outcome. Not all *lis pendens* result in a foreclosure auction, because the borrower may be able to resolve the delinquency or modify the mortgage in order to keep the home. Accordingly, we will look separately at children living in properties that sell at auctions, go to REO, or sell through arms-length sales within 12 months of the foreclosure notice.

Preliminary match rates between parcel and student enrollment data for each point in time

DOE has not yet finished matching our foreclosure data to students in the 2003-2004 and 2006-2007 school years. However, we can report on a preliminary test used to assess the feasibility of linking the student and foreclosure data sets. In that test, LP addresses were matched to student addresses in two school years (2004-05 and 2007-08), using the address data from October 31st only. This match did not adopt any of the strategies discussed below for accounting for variation in address format, and did not attempt to match on the March and June dates. Thus, this method surely undercounts the number of students whose families have faced foreclosure. But this initial, very conservative, test provides preliminary evidence that a sizable – and growing – number of New York City school children are experiencing foreclosures. Moreover, a significant share of foreclosed properties house children attending public school.

Date of foreclosure filing	N Foreclosed Properties	School year	N Students*	N (%) Students in Foreclosed Housing	N (%) of all Foreclosed Units with public school students
7/1/04-12/31/05	9,973	2004-2005	601,312	7,428 (1.2%)	3,506 (35.2%)
7/1/07-12/31/08	21,340	2007-2008	913,479	22,952 (2.5%)	9,472 (44.4%)

⁴ Note this group is identical to group a in the more conservative definition because we do not want to count students as going through a foreclosure during school year 2003-2004 if the LP was issued during the prior school year. In our analysis, however, we will still test whether foreclosures starting during a previous school year affect a child’s behavior and experiences during the current school year.

- (i) Data from 2004-05 is available only for students in grades 1-8; 2007-08 includes students in grades 1-12.
- (ii) This analysis does not exclude the students who have no usable address (address is missing, listed as a PO Box, or not in any of the five boroughs).

Other data quality issues

DOE is still in the process of matching the student addresses. Because the addresses have not been entered at the school-level in any consistent way, DOE is using several methods to account for:

- 1) Different spellings and abbreviation of thoroughfare identifiers (i.e., street can be spelled out or abbreviated as St. and refer to the same address). The programmer has identified a match key based on street number, street name, the first letter of street, boulevard, avenue, and zip code.
- 2) In New York City, many addresses contain the word East or West and these may be written out or abbreviated as E. or W. (for example East 25th Street and E. 25th Street). Here the programmer will use both in the match key, similar to 1) above
- 3) Addresses for students living in Queens are particularly problematic to match given the idiosyncrasies of addresses in this borough (for example, addresses are typically written as 123-45 42nd Street or 123 45 42nd Street, and a similar address may be 123-45 42nd Avenue, 42nd Road, or 42nd Terrace. To increase the probability of matching addresses in this borough, addresses will be standardized as 12345 42nd Street and the match key will account for Avenue, Road, Terrace, etc.

Figure 1: More Conservative Definition of Children Facing Foreclosure

Date of Student Living at Address	Date of Lis Pendens										
	June 2003	August 2003	September 2003	October 2003	November 2003	December 2003	January 2004	February 2004	March 2004	April 2004	May 2004
6/1/2003											
10/31/2003											
3/1/2004											
	Foreclosure window: if an LP is issued during the shaded window, we will assume the student's family has experienced a foreclosure.										

Figure 2: More Liberal Definition of Children Facing Foreclosure

Date of Student Living at Address	Date of Lis Pendens										
	June 2003	August 2003	September 2003	October 2003	November 2003	December 2003	January 2004	February 2004	March 2004	April 2004	May 2004
6/1/2003											
10/31/2003	Property does not sell after LP is issued and before 10/31/03										
3/1/2004					Property does not sell after LP is issued and before 3/1/04						
6/1/2004									Property does not sell after LP is issued and before 6/1/04		
Foreclosure window: if an LP is issued during the shaded window, we will assume the student's family has experienced a foreclosure.											