

NBER WORKING PAPER SERIES

FORECLOSURE EXTERNALITIES:  
SOME NEW EVIDENCE

Kristopher Gerardi  
Eric Rosenblatt  
Paul S. Willen  
Vincent Yao

Working Paper 18353  
<http://www.nber.org/papers/w18353>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
September 2012

Thanks to Chris Cunningham, Chris Foote, Scott Frame, Lauren Lambie-Hanson, Steve Ross, Joe Tracy for helpful conversations and to Stefano Giglio, Susan Wachter and John Campbell for excellent discussions at the 2012 Spring HULM, Mid-year AREUEA and NBER Summer Institute conferences respectively. The views expressed in this paper are those of the authors and not the official position of Fannie Mae, nor of any part of the Federal Reserve System, nor of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at <http://www.nber.org/papers/w18353.ack>

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Kristopher Gerardi, Eric Rosenblatt, Paul S. Willen, and Vincent Yao. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Foreclosure externalities: Some new evidence

Kristopher Gerardi, Eric Rosenblatt, Paul S. Willen, and Vincent Yao

NBER Working Paper No. 18353

September 2012

JEL No. G12,R20,R30

### **ABSTRACT**

In a recent set of influential papers, researchers have argued that residential mortgage foreclosures reduce the sale prices of nearby properties. We revisit this issue using a more robust identification strategy combined with new data that contain information on the location of properties secured by seriously delinquent mortgages and information on the condition of foreclosed properties. We find that while properties in virtually all stages of distress have statistically significant, negative effects on nearby home values, the magnitudes are economically small, peak before the distressed properties complete the foreclosure process, and go to zero about a year after the bank sells the property to a new homeowner. The estimates are very sensitive to the condition of the distressed property, with a positive correlation existing between house price growth and foreclosed properties identified as being in “above average” condition. We argue that the most plausible explanation for these results is an externality resulting from reduced investment by owners of distressed property. Our analysis shows that policies that slow the transition from delinquency to foreclosure likely exacerbate the negative effect of mortgage distress on house prices.

Kristopher Gerardi  
Federal Reserve Bank of Atlanta  
1000 Peachtree St. NE  
Atlanta, GA 30309  
kristopher.gerardi@atl.frb.org

Eric Rosenblatt  
Fannie Mae  
16715 Keats Terrace  
Derwood, MD 20855  
eric\_rosenblatt@fanniemae.com

Paul S. Willen  
Federal Reserve Bank of Boston  
600 Atlantic Avenue  
Boston, MA 02210-2204  
and NBER  
willen968@gmail.com

Vincent Yao  
Fannie Mae  
3900 Wisconsin Avenue, NW  
Washington, DC 20016-2892  
vincent\_w\_yao@fanniemae.com

# 1 Introduction

Researchers have argued that foreclosures reduce the sale prices of nearby properties. We revisit this question using a new dataset that allows us to identify and locate properties at various stages of distress, from minor delinquency all the way through the foreclosure process to lender ownership and sale to a new homeowner. Additionally, a subset of our data includes information about the condition of the foreclosed properties.

In the existing literature, researchers have typically estimated some variant of the following spatial externality regression

$$\log(P_{it}) = \alpha + \beta X_{it} + \gamma NF_{it} + \varepsilon_{it} \quad (1)$$

where  $P_{it}$  is the sale price of property  $i$  in period  $t$ ,  $X_{it}$  is a vector of controls, and  $NF_{it}$  is a measure of the number of properties that experience some type of foreclosure event within a certain distance of property  $i$  in some window around period  $t$ . There are substantial differences in the types of foreclosure events, the distances, and the time windows that previous papers have focused on, but in general, researchers have found negative estimates for  $\gamma$ , which they interpret as evidence of the existence of negative foreclosure externalities.

We estimate a spatial regression that is similar to equation (1), but with several important differences. We argue that our specification improves the identification of a true causal impact of foreclosures on prices and narrows the possible interpretations of the externality. There are three main innovations in our approach. The first is that we use multiple measures of the stock of distressed properties, whereas previous researchers have focused on a single flow. Most papers in the literature have measured the flow of properties completing the foreclosure process.<sup>1</sup> This implicitly assumes that the

---

<sup>1</sup>For example, Immergluck and Smith (2006) measured the number of transitions of property from serious delinquency into lender ownership.

foreclosure externality does not occur until the time of the foreclosure auction. In contrast, in our baseline specification, we include the number of properties with seriously delinquent mortgages (SDQs), which we define as properties owned by borrowers who have been delinquent 90 days or more on their mortgages for at least one year, the number of lender-owned properties, known in the industry as Real Estate Owned (REOs), and the number of properties recently sold by the lender. Furthermore, in variations of our baseline specification we also include the number of properties with mortgages that have been seriously delinquent for less than a year and properties with mortgages that are fewer than 90 days delinquent, which we refer to as minor delinquencies. Thus, we allow for the possibility that the foreclosure externality occurs well before the foreclosure is completed, when the borrower first becomes distressed.

The main reason to focus on stocks and not flows is that for many of the theories of why foreclosures might affect prices, it is the inventory that matters and not the flow. For example, many have argued that borrowers facing foreclosure have little reason to invest in their properties, which could generate negative externalities in the neighborhood and depress nearby home values. But the approaches used in the previous literature only roughly approximate the number of nearby properties in distress at the time of the sale.<sup>2</sup>

As we discuss below, our focus on the stock or inventory is important for policy reasons. If one interprets equation (1) causally, then flow measures can lead to erroneous inference. For example, suppose that all distressed properties exert downward pressure on prices due to investment externalities, but that equation (1) is estimated using only transitions into foreclosure. Be-

---

<sup>2</sup>For example, counting foreclosure process initiations over the last 18 months prior to a sale (as in Schuetz, Been, and Ellen (2008)), works only if foreclosure timelines do not differ substantially over time or across jurisdictions. If, for example, foreclosure timelines before the crisis rarely exceeded 18 months and after the crisis almost always did, then the measure in Schuetz, Been, and Ellen (2008) would systematically understate the growth in the stock of distressed properties.

cause foreclosure transitions in a given area are highly correlated with the number of outstanding distressed properties in the same area, one would find a significant, negative correlation between the sale price of a non-distressed property and the number of surrounding properties transitioning into foreclosure. Based on such results, one might conclude that implementing a foreclosure moratorium would increase house prices. However, such a conclusion would be wrong. Delaying transitions into foreclosure does not reduce the total number of distressed properties, which is what exerts downward pressure on prices according to the true model. Indeed, over time, delaying foreclosures without stopping transitions into delinquency would increase the total number of distressed properties and thus serve to lower prices.

Consistent with such a theory, we find that properties in all stages of distress exert downward pressure on nearby home values. Estimating a variant of equation (1), we find estimates of  $\gamma$  that are smallest in absolute value for the number of nearby minor delinquencies and largest for the number of properties with seriously delinquent mortgage borrowers that have not yet completed foreclosure proceedings. Our estimate of  $\gamma$  is slightly lower in absolute value when the lender owns the property, then falls further after the sale out of REO to an arms-length buyer, and finally reaches zero approximately one year after the REO sale.

The second innovation, which is discussed in more detail in Section 2, is the manner in which we attempt to control for unobserved heterogeneity across properties. Unobserved heterogeneity is a serious issue in this context, as it is well known that foreclosures are generated by falling house prices, so any unobserved factor that causes a decrease in house prices and thus an increase in foreclosures will lead to simultaneity bias and erroneous inference. To deal with this issue, we estimate a version of equation (1) that controls for previous sales of the same property *and* contains a set of highly geographically disaggregated fixed effects (at the census block group level). Thus, our estimates of  $\gamma$  in equation (1) reflect differences in price growth across

properties bought and sold in the same year within the same census block group (CBG). We argue that this identification strategy is largely immune to issues of reverse causality and simultaneity bias. In addition, we show that the inclusion of highly disaggregated geographic fixed effects dramatically reduces the estimated impact of nearby distressed properties on home values, suggesting that most of the previous papers in the literature, which did not employ such fixed effects, significantly overstated the magnitude of the true foreclosure externality. Our estimates of the negative impact of a nearby distressed property on the sale price of a non-distressed property are economically small in magnitude, ranging from just under 0.5 percent to just over 1.0 percent depending on the exact regression specification, the sample period, and the assumptions made about the effect of distance.

The final major innovation in the analysis is the fact that the dataset includes information on whether a seriously delinquent property is vacant and on the condition of lender-owned properties. We find that the estimate of  $\gamma$  in equation (1) is more negative for both vacant properties and lender-owned properties in “below average” condition, while the estimate of  $\gamma$  for lender-owned properties in “above average” condition is actually positive.

In Section 5 we provide an interpretation of these results. We evaluate three possible explanations: 1.) unobserved relative demand shocks that drive down prices and result in some foreclosures; 2.) foreclosures generating increased relative supply and driving down prices; 3.) an externality of reduced investment by distressed borrowers in the delinquency phase and financial institutions in the lender-ownership phase. Given the data and the limited theory, it is very difficult to establish anything conclusively. However, we argue that the weight of the evidence points to the third explanation. Both of the first two explanations require that there be distinct within-CBG micro-markets not generated by the externality from the foreclosures themselves. Given the small size of CBGs, this seems unlikely. In addition, the evidence from the regressions that incorporate information on the condition

of foreclosed properties is inconsistent with the supply explanation: a reasonable hypothesis is that foreclosed properties in above-average condition should compete more for buyers than foreclosed properties in poor condition, implying that foreclosed properties in above-average condition would have a negative impact on price rather than a positive one.

The paper proceeds as follows. Section 2 contains a discussion of the empirical approach with an extensive discussion of both the empirical model and the data. In Section 3, we report the regression results, and in Section 4 we discuss the existing literature and how this paper fits into it. We pay particular attention to a recent study by Campbell, Giglio, and Pathak (2011), which uses similar empirical methods, but adopts a different interpretation of the data that leads to differences in the reported magnitudes of foreclosure externalities. Section 5 discusses the potential interpretation of our empirical results, while the conclusion contains a discussion of the policy implications of our analysis.

## 2 Empirical Methodology and Data

We consider a sample of properties  $i \in I$ , located in geography  $g \in G$ , all sold in year  $T$  and purchased at various years  $t$  in the past. The baseline specification, shown below, is a regression of individual property price growth between years  $t$  and  $T$  on the number of nearby distressed properties in a given year  $T$  and a set of controls:

$$\log(P_{igT}/P_{igt}) = \alpha_{gt} + \beta X_{iT} + \sum_{d \in D} \gamma_d N_{iT}^d + \epsilon_{igT}. \quad (2)$$

In equation (2),  $P_{igT}/P_{igt}$  is price growth from purchase to sale,  $\alpha_{gt}$  is a full set of location  $\times$  year-of-purchase fixed effects, and  $X_{iT}$  is a vector of hedonic controls measured in year  $T$ . The variables of interest are counts of distressed property of type  $d \in D$  around property  $i$  in the year of sale. Note that  $d$  can

Table 1: Measures of Distressed Property

Measure of Distress	Distance	
	< 0.1 miles	0.1-0.25 miles
SDQ	$\gamma_{SDQ,close}$	$\gamma_{SDQ,far}$
REO	$\gamma_{REO,close}$	$\gamma_{REO,far}$
REO Sales < 1 yr	$\gamma_{SALE<1,close}$	$\gamma_{SALE<1,far}$
REO Sales 1 – 2 yrs	$\gamma_{SALE>1,close}$	$\gamma_{SALE>1,far}$

differ in both the type of distress, REO versus serious delinquency (SDQ) for example, as well as in the distance from the sale of property  $i$ . An example of  $d$  is the number of properties in REO inventory between 0.10 and 0.25 mile from the sale of property  $i$  in year  $T$ . In our main specification, we consider eight different measures of distress, displayed in Table 1.

We draw the reader’s attention to three important points about our specification. First, we use a repeat-sales approach to control for time-invariant heterogeneity across properties. Much of the previous literature has estimated hedonic models in which the dependent variable is the logarithm of the sale price in year  $T$ , and the set of control variables includes characteristics of property  $i$  in year  $T$ . The advantage of using the repeat-sales approach is that while a hedonic model usually controls for only the characteristics of the property that are contained in the tax assessor’s data base, the repeat-sales model in some sense controls for the previous sale price, which in principle captures a lot of relevant information, including everything from period detail to water views and southern exposures. It is important to stress that the repeat-sales model addresses only time-invariant characteristics of the property, and thus cannot help with reverse causality, as foreclosures are, by nature, time-varying. That said, hedonic models as they are typically implemented also use only current information about the characteristics of the property, and thus control only for time-invariant factors.<sup>3</sup>

<sup>3</sup>For example, the dataset used by Campbell, Giglio, and Pathak (2011) contains only



Because we are looking at differences in prices, in our baseline specification, we follow Harding, Rosenblatt, and Yao (2009) and define the various measures of nearby distressed properties as the *difference* in the number of properties over the repeat-sale period (the interval between year  $t$  and year  $T$ ). For example, if we consider a repeat sale of a property purchased on July 21, 2004 and sold on April 3, 2009, one measure of  $N_{iT}^d$  would be the difference in the REO inventory within 0.10 mile of the property on those two dates. This is important because it is common for a single property to go through foreclosure multiple times over the span of a few years and we would underestimate the impact of nearby distressed properties on non-distressed prices if we failed to take this into account.

Second, we include location $\times$ year-of-purchase fixed effects,  $\alpha_{gt}$ . Only a few previous papers in the literature have included any type of geographic controls, and none have included geographic fixed effects measured at the level of disaggregation used in this analysis, the CBG. For example, Campbell, Giglio, and Pathak (2011) includes a set of census tract fixed effects. Figure 1 provides an example of the breakdown of census tracts, census block groups, and census blocks for the city of Cambridge, Massachusetts, a small city of about 100,000 people that neighbors the city of Boston to the north. Cambridge is made up of 32 census tracts, and each census tract typically includes two or three census block groups.<sup>4</sup> Thus, a typical CBG is a very small geographic area, is composed of a relatively homogeneous housing stock, and has a relatively homogeneous population with respect to ethnic and economic characteristics. As we argue below, a CBG is likely smaller in geographical terms than what we typically think of as a local housing market. Note that we constrain the sample to repeat-sale pairs where the second sale occurs in the same year (2009 in our baseline specification), meaning that we need

---

information from the most recent assessor's files.

<sup>4</sup>There are over 200,000 CBGs in the United States, with each group generally containing between 600 and 3,000 people. They are subsets of census tracts, which contain between 1,500 and 8,000 people.

not include year of sale fixed effects in addition to the year of purchase fixed effects.

Third, in all of the regressions we adopt a parsimonious approach and use the unweighted number of each type of distressed property within a given radius of the non-distressed repeat sale. This assumption is common in much of the previous literature.<sup>5</sup>

Finally, in our baseline specification, we divide the distressed properties into two different distance bins. We consider an inner ring of distressed properties within 0.10 miles and an outer ring between 0.10 and 0.25 miles. The reason for including two mutually exclusive geographic areas is that there is a strong positive correlation among counts of distressed properties at different distances from a given non-distressed sale, as will become apparent in our discussion of the summary statistics below. As a result, if we are interested in the effect of distressed properties on the sale price of a property very close by, but omit the number of distressed properties farther away, our estimate will be biased upwards if the distressed properties farther away also exert a negative externality. We find that distressed properties as far away as 0.25 mile seem to exert a non-trivial negative effect on sale prices of non-distressed properties, while distressed properties at distances beyond 0.25 mile exert either no effect or an effect that is extremely small in magnitude and thus can be ignored. This finding (see our discussion of Table 7 below for more details) guides our decision to include counts of distressed properties in the second distance bin (between 0.10 and 0.25 mile) in our baseline specification.

---

<sup>5</sup>One notable exception is Campbell, Giglio, and Pathak (2011). See the discussion on page 2125 for a detailed description of their weights. However, the authors show that their results are largely unchanged by using an unweighted approach. Furthermore, since theory does not provide much guidance on the appropriate weighting scheme to use and since our distance measures are approximate, we are concerned that any inference drawn from a complicated weighting scheme may be misleading, and therefore we choose to estimate unweighted regressions.

## 2.1 Identification issues

As mentioned in the introduction, the key issue in estimating equation (2) is the possibility of reverse causality, that is to say, the possibility that the prices are driving the foreclosures and not the other way around. In a sense the use of the phrase “reverse causality” to describe the effect of prices is somewhat misleading. While the size and even existence of an effect of foreclosures on prices remains an open question, both economic theory and the data are unequivocal that prices have a strong effect on foreclosures. In this paper, we deal with this identification problem with what we describe as a “brute force” method. The combination of the repeat-sales model with the CBG $\times$ year-of-purchase fixed effects means that we are identifying  $\gamma$  in equation (2) using variation in the price appreciation of properties in the same CBG that were bought in the same year and sold in the same year. In other words, to explain a significant coefficient estimate associated with the presence of nearby distressed property, one must come up with a story about why properties within the same CBG with a higher concentration of nearby distressed property, appreciated differently from properties elsewhere in the same CBG over the same time interval with a smaller concentration of nearby distressed property. For example, the fact that properties on the main street in a given CBG are in higher demand and thus more valuable than properties off the main street with fewer nearby foreclosures in the same CBG would not generate a significant negative estimate of  $\gamma$ . The difference in demand stemming from this within-CBG location difference is present at both times  $t$  and  $T$ , and thus is subsumed by expressing the dependent variable as the difference in sale prices. Rather, one would need to tell a story about why prices *fell* between times  $t$  and  $T$  in one area of a CBG relative to another area in the same CBG.

Using within-CBG variation to identify the effect of nearby distressed property on price appreciation also significantly alleviates concerns about reverse causality. Figure 2 provides an example of how using variation across

geographies could cause one to mistakenly conclude that there was a causal effect of nearby foreclosures on prices when the true causal effect actually went from prices to foreclosures. In the example, we assume that we have data on foreclosures and prices in two separate geographic areas, tract A and tract B, as shown in the top panel. We assume that price appreciation is constant for all properties within the same geography, but that price appreciation is significantly lower in tract B than in tract A. Within each tract, foreclosures are randomly located so there is, by construction, no causal effect of nearby foreclosures on prices. Thus, separate within-tract regressions of sale prices on the number of foreclosures within some radius would correctly yield a  $\gamma$  of zero. In the bottom panel of Figure 2 we look at the same data but ignore the geographic differences. A regression of sale prices on the number of nearby foreclosures now incorrectly yields a negative relationship between price growth and foreclosures. This simple example illustrates how reverse causality can confound identification, and in the analysis below, we show that the estimates of  $\gamma$  in equation (2) are indeed quite sensitive to the inclusion of and the aggregation level of geographic fixed effects.

## 2.2 Data

Our sample of repeat sales includes all pairs of non-distressed transactions on single-family residential properties in the 15 largest metropolitan statistical areas (MSA), pulled from public records purchased from a national data aggregator. The sample is restricted to include transactions in which the first sale in the repeat-sale pair took place after 2001 and the second sale took place between 2006 and 2010, allowing us to study both the pre-crisis and post-crisis periods.<sup>6</sup> We exclude addresses that cannot be geo-coded and transactions for which recorded prices or dates are missing, zero, or located

---

<sup>6</sup>The previous literature has, for the most part, used pre-crisis sample periods. The exception is Campbell, Giglio, and Pathak (2011), which uses data between 1987 and the first quarter of 2009, and thus captures a good portion of the crisis period.

in a thin market, which we define as a CBG in which there were fewer than five sales in a year. The final sample contains 958,513 repeat-sale pairs in 15 MSAs, and 16,932 CBGs, as reported in Panel (A) of Table 2.<sup>7</sup> The bottom panel of Table 2 reports the distribution of observations in the repeat sales sample by the year of purchase and year of sale, respectively. There are several notable patterns in the table. The sample of repeat sales gets smaller over time as national sales volumes fall. In 2006 and 2007, the modal sale occurred two years after purchase, increasing to three years in 2008, four years in 2009, and five years in 2010. The MSAs with the most observations are Phoenix, AZ, Washington, D.C., and Riverside, CA, which account for 16 percent, 13 percent, and 10 percent of the sample, respectively. Table 3 shows that the repeat-sales sample includes enormous variation in returns, which is not surprising, given that the dataset includes properties purchased in 2001 and sold in 2006 and also properties purchased in 2006 and sold in 2010. The public records data also contain information on basic property characteristics over time, including house size, lot size, property age, and number of bedrooms. These variables are summarized in Table 3.

Using the public records, we can identify the date of the foreclosure deed,<sup>8</sup> when the lender records transfer of ownership from the borrower, and the REO sale date, when an arms-length buyer takes ownership of the property. Using these flows, we can compute foreclosure inventory in a location at any point in time. The final sample contains 1.04 million foreclosure deeds, which we refer to as the REO inventory throughout the paper, and 1.15 million REO

---

<sup>7</sup>Another important difference with much of the previous literature is the national representativeness of our data. Many previous studies focus on a single state or even a single MSA. For example, Campbell, Giglio, and Pathak (2011) uses data from the state of Massachusetts, Immergluck and Smith (2006) uses data from Chicago, and Schuetz, Been, and Ellen (2008) uses data from New York City. Harding, Rosenblatt, and Yao (2011), which uses data from seven MSAs, is probably the most nationally representative study.

<sup>8</sup>The foreclosure deed corresponds either to the transfer of the property to the lender at auction, or if the auction is successful, to the transfer of the property directly to another arms-length buyer. The latter event is significantly less likely to occur than the former, but we are able to distinguish between the two events in the data.

sales.

To identify seriously delinquent properties, we use two methods. Our main approach is to use proprietary data from a large national mortgage insurer (the proprietary data mentioned in the introduction), which contain all of the information in the public records plus a detailed payment history. This dataset allows us to identify the first month in which a delinquent borrower enters serious delinquency (SDQ), which we define to be 90 days delinquent (typically three missed payments). An SDQ corresponds to the entire period in which the borrower is seriously delinquent before the foreclosure auction, and thus it covers both the time before the foreclosure process is initiated on a seriously delinquent borrower, and the time between the start of the foreclosure process and the end of the process (the auction).

The data also allow us to identify the cumulative depth of delinquency at any point in time. Our dataset contains 1.12 million SDQs. Because the proprietary dataset does not cover the universe of all homes, we augment it with data from a nationally representative loan-level dataset (the LPS data). With the more representative dataset, we calculate, for each state, the distribution of the number of months that it takes for a mortgage to transition from serious delinquency to foreclosure completion (the foreclosure auction). We then take the 25th percentile of these distributions and combine them with the information from the public records database on the date of the foreclosure auction to impute SDQ intervals. For example, the 25th percentile for California is four months. Thus, for each of the REO properties located in a California MSA in our sample, we assign an SDQ interval corresponding to the four months before the foreclosure auction dates. To be conservative, we use the 25th percentile as opposed to the median or average, as this means that 75 percent of foreclosures in California had a serious delinquency spell that lasted for more than four months. We call this variable “infilled” SDQs. This provides 726,547 additional SDQs. We then combine our infilled SDQs with the SDQs obtained from the proprietary mortgage database to produce

a more encompassing SDQ measure.

For our analysis, we divide SDQs into “long SDQs” and “short SDQs,” depending on whether the borrower has been delinquent for more than one year or not. For some regressions, we also look at “minor DQs,” which we define to be delinquencies of 60 days or fewer. In addition, we have collected property-level vacancy data from the U.S. Postal Service. Postmen classify each property as “occupied” or “vacant” each month based on whether or not mail is being accepted at each property. These data are currently being used by mortgage servicers to identify loss mitigation opportunities. As we mention above, we also have some information on the condition of lender-owned properties. These data come from REO appraisals ordered by lenders immediately after a property completes the foreclosure process. The full residential appraisal form provides a detailed description of the current condition of the property that typically includes a standard rating index with values of “great,” “good,” “average,” “fair,” and “poor.” For the appraisals that do not include the rating index, but do include detailed descriptions, we text-mined all of the keywords in the description and sorted them into the standard condition rating according to the Marshal and Swift *Residential Cost Handbook* used by appraisers. We characterize “great” and “good” as “above average,” and “fair” and “poor” as “below average.”

Panel A of Table 4 shows that two thirds of repeat sales had no distressed properties nearby. Panel B considers differences in the number of nearby distressed properties between the second and first sale in the repeat-sale pair, and shows that a nontrivial number of repeat sales had *fewer* distressed properties nearby at the time of the second sale, with roughly 5 percent of sales occurring near properties with lower REO inventory and fewer sales of REO in the year preceding the sale. Panel C of Table 4 shows that, not surprisingly, the incidence of sales with distressed properties nearby has increased significantly over time. Most dramatically, the proportion of properties with long SDQs nearby rose from less than 2 percent in 2006 to more than 30

percent in 2010, reflecting both the increased hazard that borrowers transition into serious delinquency and delays in the foreclosure process. By 2010, more than half the sales in our sample occurred with at least one form of distressed property nearby.

Finally, panel D of Table 4 displays correlations between our stock measures of distressed property and flow measures for both the close distance (within 0.10 mile) and the far distance (between 0.10 and 0.25 mile). Comparing the stock and flow measures of close distressed properties (first four rows and first four columns), it is apparent that they are all positively correlated. However, no two measures have a correlation higher than 0.50, a fact that emphasizes the importance of distinguishing between stocks and flows. We see similar patterns for the correlations among the different types of distressed properties at the far distance (last four rows and last four columns). However, the positive correlations between the number of the same types of distressed properties at different distances are very strong. For example, the correlation between REO inventory at the close distance and REO inventory at the far distance is 0.65, while the corresponding correlation for REO sales in the year before the repeat sale is 0.73. As we discuss above, these correlations emphasize the need to control for distressed property at longer distances in the estimation of equation (2), which we do in all of the regressions reported below.

### 3 Results

Table 5 shows results from our baseline specification. The right-hand-side variables of interest are nearby long SDQs from the proprietary data and three measures from the public records: nearby REO inventory, the number of nearby REOs sold one year prior to the non-distressed sale, and the number of nearby REOs sold one to two years prior to the non-distressed sale. For each variable, we measure the difference over the repeat sales in the number



within 0.10 mile (about 530 feet), which we label “close” and the number between 0.10 and 0.25 mile, which we label “far.” Despite the fact that we are using repeat sales, we control for the possibility that there is systematic variation in price growth across different types of properties by including the characteristics of the property from tax assessment data.<sup>9</sup> In addition, to control for the possibility that prices fell more in more dense areas within a given CBG, we include the number of properties within 0.10 mile in our baseline specification.

Column (1) of Panel A in Table 5 displays the results of our basic specification on 2009 data (repeat sales for which the second sale occurred in 2009). The estimation results for the variables of interest at the close distance show a basic pattern that is replicated in all of the subsequent regressions: the coefficient estimates associated with the first three stages of the foreclosure process,  $\gamma_{SDQ,close}$ ,  $\gamma_{REO,close}$ , and  $\gamma_{SALE<1,close}$ , have roughly similar magnitudes, with nearby serious delinquencies having the largest magnitude (in absolute value), and recent nearby REO sales having the smallest. The exception is nearby REO sales that occurred more than one year in the past, which are not negatively correlated with price growth. Columns (2) and (3) of Panel A show that the controls have only small effects. Another pattern that emerges from the table is that the measures of close distressed property have a slightly more negative effect on price growth than the measures of far distressed property. For example, an additional long SDQ within 0.10 mile is estimated to decrease price growth by 1.2 percent while the effect of an additional long SDQ between 0.10 and 0.25 mile is -0.8 percent. Most previous

---

<sup>9</sup>If the logarithm of the price of a property is a linear function of time invariant characteristics then, as Harding, Rosenblatt, and Yao (2009) shows, taking the difference over the repeat sales will cause the characteristics to cancel out of the equation; therefore one does not need to control for them in the repeat-sales specification. In other words, this implicitly assumes that price growth is not a function of the time-invariant characteristics of a property. However, because of preference changes, it may be the case that properties with different characteristics appreciate at different rates. For example, if homes with multiple bathrooms or with granite countertops become more sought after over time, then we would expect those properties to appreciate at a higher rate, all else being equal.

studies have also found that the effect of foreclosures on nearby home values decreases with distance. We discuss this literature in more detail in Section 4 below. The corresponding magnitudes for the REO inventory and REO sale measures are very similar, although the difference between the close and far estimates for REO inventory is very close to zero.

Panel B in Table 5 shows how the coefficient estimates in the baseline specification change over time. In the panel we estimate the baseline specification of equation (2) separately for repeat-sale pairs in which the second sale took place in each year between 2006 and 2010. The coefficient estimates associated with long SDQs and REOs sold one to two years within 0.10 mile before the non-distressed sale,  $\gamma_{SDQ,close}$  and  $\gamma_{SALE>1,close}$ , are the largest and smallest, respectively, in absolute value for all five years. The estimated impact of long SDQs within 0.10 mile and between 0.10 and 0.25 mile is considerably larger in 2006 and 2007 than in later years, but this pattern does not seem to hold for our other measures of distressed property. The coefficient estimates associated with close REO inventory and REOs sold in the year before the non-distressed sale are quite stable, with the former variable having a slightly smaller impact than the long SDQ variable for the 2009 and 2010 samples, and a significantly smaller impact for the 2006, 2007, and 2008 samples. The estimated coefficient associated with the close REO sale variable grows over the sample but is consistently smaller in absolute value than the estimated coefficients associated with both long SDQs and REO inventory.<sup>10</sup>

---

<sup>10</sup>There is some evidence from the literature that the effect of nearby foreclosures on prices is non-linear, specifically that it is diminishing in the number of nearby foreclosures. In unreported regressions, which are available from the authors upon request, we explored this using a more flexible specification, in which we specified the number of nearby distressed properties as second and third order polynomials and included a series of indicator variables for each specific value. Consistent with the findings from the previous literature, we did find evidence of nonlinearities, as the effect of nearby distressed properties on prices is diminishing in the number of distressed properties. However, all of the results discussed in this section are robust to this more flexible specification, and thus for space considerations we chose to report the simpler linear specifications.

The inclusion of  $\text{CBG} \times \text{year-of-purchase}$  fixed effects in the baseline specification plays a significant role in the estimation results. Panel C in Table 5 shows that the coefficient estimates associated with nearby distressed property become much stronger with more aggregated  $\text{geographic} \times \text{year-of-purchase}$  fixed effects, as the coefficient estimates associated with nearby long SDQs, REO inventory, and REO sales in the previous year approximately double in absolute value when we move from a specification that includes  $\text{CBG} \times \text{year-of-purchase}$  fixed effects to a specification that completely excludes  $\text{geographic} \times \text{year}$  effects. This confirms the intuition from the example that we discussed in Figure 2 in Section 2.1. The results in Panel C of Table 5 show that it is across-census tract and across-MSA variation in foreclosure density that is driving much of the observed negative correlation of nearby foreclosures and prices at the national level. The estimates for each of the three variables of interest increase substantially in absolute value when we substitute  $\text{county} \times \text{year}$  effects for  $\text{census tract} \times \text{year}$  effects and increase again when we move from  $\text{MSA} \times \text{year}$  effects to eliminating  $\text{geographic} \times \text{year}$  effects altogether. The last two columns in Panel C show that the results are little changed when substituting  $\text{CBG} \times \text{year-of-purchase}$  fixed effects for  $\text{census tract} \times \text{year}$  effects, which suggests that the census tract is a sufficiently small geography to deal with reverse causality and simultaneity bias in this context. With the exception of Campbell, Giglio, and Pathak (2011), who include census tract geographic controls, all previous attempts to estimate  $\gamma$  omit narrow geographic controls, and thus most likely significantly overestimate the effect.

In Table 6, we exploit information about the vacancy status of distressed property and the condition of REO property. As discussed in the previous section, for a subset of SDQs, we have information about the vacancy status of the properties, and for a subset of the REO properties we have information about condition. Since the vacancy data, in particular, are only well populated beginning in 2010, we focus on that year. The results show that the

coefficient estimate associated with vacant SDQ property is approximately 50 percent larger in absolute value than the coefficient estimate associated with occupied property (-0.011 versus -0.006). But, perhaps the more significant results apply to the condition of the REO inventory. According to Table 6, the only significantly negative coefficient estimates are associated with REO in below average condition and with REO for which we do not have condition information. The fact that the estimate associated with the missing category is significantly negative likely reflects the fact that most REO is in below average condition. It is also worth noting that the estimated coefficient associated with above average REO is significantly positive, suggesting that nearby REO in good condition actually *increases* the sale price of non-distressed properties.

In Table 7 we look at the impact of distressed properties even farther away from the repeat-sale observations, by augmenting the baseline specification for 2009 data with two additional rings of property: 0.25–0.50 mile away and 0.50–1.0 mile away. As previous researchers have found, the negative effect of nearby distressed property drops off very quickly with distance, as the coefficient estimates associated with distressed properties in the third and fourth rings are very close to zero. As we discussed above, the results in Table 7 motivate our decision to include only the 0.10–0.25-mile ring in our baseline specification, as distressed properties farther than 0.25 miles do not exert an economically meaningful effect on non-distressed home sales.

In Table 8 we allow for the possibility that properties with moderately delinquent mortgages also exert a negative externality. The table reports estimation results in which we distinguish between short and long SDQs. The results in the first two columns of the table suggest that both short and long SDQs have similar effects on nearby non-distressed sale prices. In column (3) we distinguish between minor DQs, short SDQs, and long SDQs. Minor DQs also have a negative estimated effect on nearby non-distressed sales, which is similar in magnitude to the effect of short SDQs. When we

distinguish between minor DQs and SDQs, we see that a relatively large difference emerges between the effect of long SDQs and short SDQs/minor DQs, with long SDQs having a much larger negative impact on non-distressed price appreciation.

In addition, we estimated a series of regressions to address specification issues. At least two issues merit special attention. The first is a potential omitted variable problem since our coverage of SDQs is only partial, because the proprietary mortgage data do not cover the entire mortgage market. Since non-proprietary SDQs are likely to be correlated with both proprietary SDQs and all other measures of distress, including the large number of non-proprietary REOs from our public records database, the omitted variables here could potentially affect all of the estimates of interest. To address this problem, we used the infill method, which we discuss in detail above, to construct a dataset of non-proprietary SDQs. Comparing columns (5) and (6) of Table 8 illustrates that using this much broader measure of SDQs makes little difference. Overall, the coefficient estimate associated with SDQ decreases slightly in absolute value; the other coefficient estimates decrease slightly as well.

The second specification issue is the choice of a repeat-sales specification rather than a hedonic approach. In Table 8, we consider two alternative approaches to estimating  $\gamma$ . In column (8), we show results from a standard hedonic model; these results are similar to the results from the repeat-sales specification. We find slightly larger coefficient estimates (in absolute value) for all of the distressed property variables. As an alternative, we consider a model, displayed in column (7) of Table 8, that lies part way between the two specifications, in which we use the repeat-sales specification but include on the right-hand-side the *level* of distressed property at the time of the second sale instead of the difference between repeat sales. This hybrid model also generates similar results, suggesting that specifying the distressed property measures as first differences rather than as levels does not make a substantive

difference.

## 4 Comparison with earlier work

This paper builds on an extensive existing literature on externalities in housing markets. Narrowly, there is a series of papers that estimate equation (1), starting with Immergluck and Smith (2006) and including: Schuetz, Been, and Ellen (2008); Rogers and Winter (2009); Harding, Rosenblatt, and Yao (2009); Lin, Rosenblatt, and Yao (2009); and Campbell, Giglio, and Pathak (2011). All of these papers use flow measures of foreclosure-related distress as the right-hand-side variables of interest. More broadly, a much older literature has estimated almost exactly the same hedonic regressions, but with other events not related to foreclosure that might affect local house values. We begin this section with a detailed discussion of the recent literature, as it is more related to our current analysis, and then provide a brief discussion of the older literature.

Although previous studies have used the repeat-sales specification and controlled for geography at a relatively disaggregated level, no analysis has done both at the same time. Harding, Rosenblatt, and Yao (2011) is the only paper to our knowledge to estimate equation (1) using a repeat-sales specification. It estimates separate regressions by MSA but does not control for geography within the MSA. This effectively means that it is comparing price growth and nearby foreclosures for non-distressed repeat sales across entire MSAs, and thus its estimates are prone to the same identification issues that we discussed above in the context of Figure 2. There are strong within-MSA patterns in price growth, with much sharper price declines in poorer neighborhoods and locations further from the city center, and our results below confirm that more disaggregated geographic controls generate a major reduction in the estimate of  $\gamma$ . Campbell, Giglio, and Pathak (2011) (hereafter CGP), to our knowledge, is the only study in the literature that

includes a disaggregated set of geographic controls. CGP uses a hedonic model and includes census-tract $\times$ year controls, which, as we saw from Figure 1, are slightly more aggregated than the CBG controls that we employ in this paper. However, as we have shown, the difference in estimates of  $\gamma$  using CBG versus census tract controls is small, which suggests that the census tract is a sufficiently small geography to eliminate the influence of unobserved heterogeneity in the estimation of equation (1). The other studies mentioned above all use hedonic models with either no disaggregated geographic controls or fairly broad ones.

While CGP does include relatively disaggregated fixed effects in its hedonic specification, the fixed effects are not the focus of its identification strategy. Instead, CGP focuses on a difference-in-differences approach. First, the authors of CGP consider two  $\gamma$ s: a  $\gamma_B$  associated with foreclosure completions that occur in the year prior to a nearby non-distressed sale and a  $\gamma_A$  associated with foreclosure completions that occur in the year after a sale. They argue that  $\gamma_A$  represents the causal effect of prices on foreclosures, writing that, “To the extent that house prices drive foreclosures, low prices should precede foreclosures rather than vice versa,” and that  $\gamma_B - \gamma_A$  therefore represents the causal effect of foreclosures on prices. In addition, CGP takes another set of differences by subtracting  $\gamma_B - \gamma_A$  estimated for a ring 0.10–0.25 mile away from the sale from  $\gamma_B - \gamma_A$  estimated for foreclosure within 0.10 mile.

Superficially, the approach in CGP appears to differ sharply from ours but, in practice, the approaches are quite similar. Following CGP, we could interpret  $\gamma_{SDQ,close}$  as the effect of prices on foreclosures and  $\gamma_{REO,close}$  as the combined effect of prices on foreclosure and foreclosures on prices and then report

$$\gamma_{REO,close} - \gamma_{SDQ,close} - (\gamma_{REO,far} - \gamma_{SDQ,far}).$$

This is because foreclosure auctions that occur after the non-distressed sale are most likely in a state of SDQ before the time of the sale. Thus, our

$\gamma_{SDQ,close}$  roughly corresponds to the  $\gamma_A$  in CGP.

But in contrast to CGP we choose to report the individual coefficients and to interpret them separately. To see why, suppose that the true mechanism for the foreclosure externality works through reduced investment. As we discuss in Section 5.2 below, theory says we should expect that seriously delinquent homeowners will underinvest, generating negative externalities for neighboring properties. Using the CGP approach, we would mistakenly identify that underinvestment as the effect of prices on foreclosures and conclude that there was no effect of foreclosures on house prices prior to the foreclosure and a severely attenuated effect after. Therefore, we believe that the coefficient estimates themselves,  $\gamma_{REO,close}$  and  $\gamma_{SDQ,close}$  (or  $\gamma_B$  and  $\gamma_A$ ), rather than the difference, is the relevant estimate of the externality.

Similarly, unlike CGP, we do not focus on the difference between the close and far coefficient estimates. The reason for this is that it is unclear whether or not distressed properties between 0.10 and 0.25 mile away exert a true causal impact on prices. If they do not, then taking this difference as CGP do would be correct, as any estimated effect of distressed properties on prices within this distance would be capturing simultaneity bias. However, if distressed properties within this distance do in fact have a causal impact on prices, then taking the difference would yield an underestimate of the true causal impact. Ideally, theory would provide some guidance on this issue, but unfortunately it does not. To put it another way, there is no *a priori* reason to think that a foreclosure 0.11 mile away has no effect on prices but a foreclosure 0.09 mile away does. But, depending on the causal mechanism in play, there may be good reasons to expect the effect of foreclosures on prices to dissipate with distance. For example, if distressed properties impact prices because of an investment externality, we would certainly expect the magnitude of such an effect to decrease with distance. However, the distance at which the effect goes to zero is not clear, and may not be 0.10 mile, which is what CGP implicitly assumes by taking the difference. For this reason, we



include and report counts of distressed property at both distance bins, but do not take the difference in the respective coefficient estimates.

We find that taking CGP's difference-in-differences estimator seriously with our data implies that a foreclosure has no economically meaningful effect on prices, as the  $\gamma$ s associated with SDQ and REO properties are roughly the same. CGP's authors do find a significant effect, but the difference between our findings and their headline number results from the inclusion of condominiums in their main sample. For the single-family residential properties that we (and most previous papers in the literature) focus on, their point estimate is close to zero and statistically insignificant.<sup>11</sup> Taken literally, the conclusion of the CGP analysis would be that foreclosures of single-family properties have no effect on the prices of other single-family properties. But a more plausible interpretation is that CGP's authors are over-differencing and, as a result, are underestimating the impact of foreclosures on nearby non-distressed home values. CGP does find and report significant  $\gamma$  coefficients associated with foreclosures that occurred both before and after the non-distressed sale, so that a more plausible interpretation of its empirical results is that  $\gamma_A$  measures foreclosure externalities that occur before the foreclosure is completed and not the causal effect of prices on foreclosures.

Since the publication of CGP, other researchers have adopted the difference-in-differences approach. Hartley (2011) looks at foreclosures in Chicago and estimates different  $\gamma$ s by structure type. He finds that while the  $\gamma$  associated with a single-family foreclosure on a single-family sale is significantly negative, the  $\gamma$  associated with a multi-family foreclosure on a single-family sale is not. Hartley interprets this as evidence that the supply of property rather than investment externalities drives the discount. However, while multifamily  $\gamma$  is statistically insignificant, it is economically large and positive, raising questions about the interpretation of the coefficient. Further, Hartley finds effects only within 0.05 mile and no effect between 0.05 and 0.10 mile, contra-

---

<sup>11</sup>See Table A-19 of the Internet Appendix to Campbell, Giglio, and Pathak (2011).

dicting most previous research, including CGP. Anenberg and Kung (2012) look at sales of single-family properties near San Francisco over the period 2007–2009 and augments CGP by including information about listings. Its authors argue that there is no foreclosure externality effect because they find that the  $\gamma$  of foreclosures that occur prior to a sale is zero. In this respect, their results contradict CGP, our findings and all of the previous research cited above.

As we noted above in the introduction, two important differences between our specification and all previous work are the use of stock measures of distressed property rather than flows, and the focus on distressed properties before rather than after the foreclosure completion. For example, Immergluck and Smith (2006) counts foreclosure deeds in the two years prior to the sale of non-distressed property; Harding, Rosenblatt, and Yao (2009) constructs a series of measures of foreclosure deeds in three month intervals before and after the sale; and CGP counts all properties for which foreclosure proceedings have been completed. By focusing on properties that have completed the foreclosure process, these papers implicitly assume that the externality does not occur until the foreclosure auction takes place. As we have seen, this is clearly not the case.<sup>12</sup>

To see this more clearly, consider a specification that includes the number of foreclosure completions that take place one year before the non-distressed sale of interest. Effectively, such a specification assumes that a property that was foreclosed on more than one year in the past plays absolutely no role whatsoever in the pricing of a nearby property. One might argue that exactly the opposite is true: the properties that produce the most blight, and that may be most likely to adversely impact surrounding values are the properties that lenders cannot sell. To make matters worse, the potential bias introduced by measuring flows instead of stocks is likely not constant

---

<sup>12</sup>An exception is Schuetz, Been, and Ellen (2008), which counts the number of foreclosure initiations, known as *lis pendens* filings in New York, in the 18 months prior to a non-distressed sale.

over time or across locations. Foreclosure timelines differ widely across states and have slowed considerably through the recent boom/bust cycle, especially in states that require judicial review.<sup>13</sup>

Until the mid-2000s, studies that estimated hedonic price regressions similar to equation (1) largely ignored foreclosures because, up to that point, foreclosures were not a major issue. Two topics of focus in the early literature were the presence of sex offenders and subsidized housing programs.<sup>14</sup> We focus on the latter because it is, in fact closely related to the topic of this paper. Many early studies using aggregated data attempted to calculate whether subsidized housing raised house prices. Galster, Tatian, and Smith (1999) developed a methodology to use transactions-level data to measure the impact of Section 8 housing<sup>15</sup> on the sale prices of neighboring properties. Its authors compared the sale prices of properties within 500 feet of a Section 8 site before and after the site transitioned to Section 8 and assumed that the difference in sale prices measures the treatment effect. Many studies in the literature subsequently used this methodology, which is very close to the strategy used by CGP that we discussed above. The most relevant to our analysis is Schwartz et al. (2006), which used both highly disaggregated location variables, and, for one specification, used the repeat-sales method rather than a hedonic specification to control for property characteristics.

All of the cited papers used similar methods, regressing log price or price growth on some measure of distressed property within a given radius. The only paper to deviate substantially is Rossi-Hansberg, Sarte, and Owens III (2010), which estimated a version of equation (1) to measure the effects of a program in Richmond, VA to subsidize investment in properties in disadvantaged neighborhoods. Its authors diverged from the literature by using

---

<sup>13</sup> See Gerardi, Lambie-Hanson, and Willen (2011) for a discussion of foreclosure timelines across states and over time.

<sup>14</sup>Linden and Rockoff (2008) estimate a version of equation (1) using data on the registration of sex offenders at particular addresses.

<sup>15</sup>Section 8 refers to the eighth section of the Housing Act of 1937, which authorized payments of rental housing assistance to landlords on behalf of low-income households.

a semiparametric approach to estimate a pricing surface for all locations in their subject area and then looked at how the investment affected the surface.

A key difference between the investment externality literature and the foreclosure externality literature is the permanence of the effect. With some possible exceptions, serious delinquency and REO status are temporary states, whereas investment is longer lived. Thus, the distinction between stocks and flows is likely less important for the investment externality question than for the foreclosure externality question.

## 5 Interpretation

In Section 3, we established some empirical facts. Houses that sell very close to all forms of distressed property appear to do so at slightly lower prices than otherwise similar properties in the same CBG that sell without the presence of nearby distressed properties. The effect appears when the borrower becomes seriously delinquent on his mortgage and disappears one year after the lender sells the foreclosed property to a new homeowner in an arms-length transaction. What can explain these empirical observations? If our identification strategy, which uses a repeat-sale specification with disaggregated geographic fixed effects, is truly successful in dealing with reverse causality and simultaneity bias, then we argue that the likely explanation for the empirical evidence in this paper is what we refer to as an “investment externality” effect. This refers to the tendency of financially distressed borrowers and lenders that do not derive consumption services from foreclosed property to underinvest in property maintenance, leading to physical deterioration of the property and, in turn, causing a reduction in the value of nearby property to potential buyers. On the other hand, if our identification strategy fails to deal successfully with the severe econometric issues that we discussed above, then two alternative explanations of the estimation results are plausible: demand shocks and supply shocks. We begin by dis-

cussing these two alternative explanations, and then turn to a discussion of investment externalities.

## 5.1 Supply and demand effects

The idea that a fall in demand leads to a fall in prices and an increase in foreclosures has strong support in theory and in the data. Default makes sense for a borrower only if he is in a position of negative equity, that is, if his mortgage debt exceeds the value of his home, and it is difficult to have negative equity without falling prices. More broadly, models in which households face only limited opportunities to take out unsecured loans yield a relationship between the extent of negative equity and the probability of foreclosure.<sup>16</sup> Studies using micro-level data confirm the theory, showing a strong relationship between equity and foreclosure, and Gerardi, Shapiro, and Willen (2009) shows that the estimated relationship is driven by within-town, within-time variation in equity and therefore is not attributable to causality running from foreclosures to prices. How does this relate to the current paper? All else being equal, a negative demand shock in location A relative to location B would lead to a fall in prices in location A and a concomitant increase in negative equity and foreclosures.

The fact that the demand theory *could* explain the observed facts does not necessarily mean that it *does* explain them. One important fact does provide some support for the demand story and that is the result in Table 8 in which we found that  $\gamma$  is significantly negative even for minor delinquencies. One could argue that it is unlikely for significant property depreciation to occur in situations in which borrowers have only missed one or two mortgage payments, and thus, this empirical finding suggests that the estimates of  $\gamma$  may be explained by reverse causality. However, it could also be the case that a borrower who has only missed a few payments may have been in financial duress for quite some time, in which case the lack of investment in property

---

<sup>16</sup>See Gerardi, Shapiro, and Willen (2007).

maintenance might be a plausible explanation for the negative coefficient estimate associated with minor DQs.

But two additional empirical facts are inconsistent with the demand explanation. The first is the fact that the estimated  $\gamma$  is so much smaller for nearby properties sold out of REO more than a year in the past. While it is without doubt possible to construct a model in which demand recovers after the foreclosure, it seems implausible. The second broader issue here is that we are looking at variation within CBGs and any demand story (and any supply story, as well, as we argue below) must operate at the market level. For the demand theory to make sense, we would have to believe that there are distinct submarkets within a census block group, a geographic area with a “target size,” according to the Census, of around 1500 people.<sup>17</sup> Recall that since we are looking at repeat sales, some shock would have to happen at a given time and impact one part of the CBG more than another. For example, one end of a CBG may have ocean views, but to explain our results, it would have to be the case that demand for ocean views rose or fell more over the relevant time period than the demand for properties in the CBG without ocean views.

A popular alternative to the demand theory is the supply theory, which posits that a foreclosure increases the supply of property on the market and drives down prices. Unlike the demand theory, the supply theory is not an obvious consequence of standard economic theory. Normally, when we price long-lived assets like houses, we define supply as all of the assets that exist, not just the ones currently for sale. Foreclosures do not change the number of houses or the quantity of land in a market, so standard models would not predict any effect on prices. Market frictions obviously could change this. When shopping for a home, buyers generally restrict their attention to properties listed for sale, and an increase in the number of options gives the buyer more bargaining power. One might then argue that we should focus our

---

<sup>17</sup>[http://www.census.gov/geo/www/cob/bg\\_metadata.html](http://www.census.gov/geo/www/cob/bg_metadata.html).

attention on the listing of the foreclosed property to identify supply effects. However, both buyers and especially sellers know about foreclosures before they transition to REO and are listed by the lender, since foreclosures are, by law, public information. Therefore, we cannot assume that the foreclosure would not affect the listing behavior of other sellers who might, for example, list prior to the foreclosure’s making the transition to REO, in order to exploit the temporarily low supply. In a sense, with forward-looking households, the question would be when the news that the property will go on the market arrives. But, if buyers and sellers are that forward-looking, then the whole short-run frictions story has less merit. To make matters more confusing, a troubled borrower facing foreclosure cannot sell his or her home—that is the reason he or she is facing foreclosure—so, according to the bargaining story above, serious delinquency should drive down supply and *increase* prices.

Our view is that the supply theory is not a very good candidate to explain the estimation results in this paper. The supply story could potentially explain why prices rise after the REO sale: with the property now off the market, prices recover to their pre-delinquency level. However, other facts from the data work against it. The first problem is the evidence on condition. If we thought foreclosed properties were driving down prices by competing with non-distressed sales, then we would expect, at the very least, that the properties in above average condition would have the same effect as properties in below average condition and, indeed, we might even expect the above-average properties to generate even more competition. The second problem with the supply story is the same one that, in our opinion, dooms the demand story: the fact that we are identifying  $\gamma$  using within-CBG variation. With that said, it is important to stress that while we doubt that variation in supply explains the estimation results, because of our empirical design, we are unable to say much about the potential existence of supply effects from distressed properties. This is because, in all likelihood, any supply effect that operates at the level of a local housing market is washed away by construction

with the inclusion of highly disaggregated geographic fixed effects in the regression specification.

## 5.2 Investment externalities

The third candidate explanation for the results in Section 3 is that foreclosures lead to an investment externality. According to this theory, neither delinquent borrowers nor lenders have an incentive to maintain the property properly, which leads to physical deterioration of the property that some have labeled a “disamenity,” and this, in turn, reduces the value of nearby property to potential buyers. The investment disincentive is arguably present both during the SDQ period and when the property is REO. There are two reasons why seriously delinquent borrowers are unlikely to maintain properties. The first is that many of them have suffered cash-flow-depleting life events and discount future consumption heavily relative to current consumption. Effectively, this raises the hurdle rate on any investment in the property. The second problem is that many seriously delinquent borrowers expect to lose their homes and, therefore, the long-term benefits of any investment would accrue to the ultimate owner of the property—the lender.

After the foreclosure, there are good reasons to expect underinvestment as well. Narrowly, the lender does not obtain any consumption benefit from investment. More broadly, the scale of residential lending leads to an information problem. The agent managing the property rarely has a large ownership stake in the property. This is obviously true when the agent is acting on behalf of a securitization trust, but it is even true when the agent works for a bank that owns the property. The issue is that the amount of profitable investment in the property is a matter of discretion, and the owner of the property cannot be sure whether the manager has other incentives. As with all standard asymmetric information problems, the result would be a failure to exploit profitable gains from trade, in this case, investment in the property. Another way to put this is that the optimal mechanism for



investment in single-family residential real estate is to sell the property to a small-scale investor who internalizes the costs and benefits.

Evidence from the existing literature supports the existence of an investment externality. First, there is ample evidence that foreclosed properties suffer from underinvestment. We will focus our discussion on studies that use data from the current crisis period, but there is an older literature that emerged well before the most recent housing market boom and bust, which we will not cover. Interested readers can find a summary of this older literature in Frame (2010).

Pennington-Cross (2006) uses a variant of the repeat-sales methodology on nationally representative data from 1995 to 1999 and finds that the price of a distressed property appreciates significantly less than the prevailing metropolitan area price index in which the property is located. The average difference is 22 percent, but it is highly sensitive to both the length of time that a property spends in the foreclosure process and the length of time that it remains in REO, as well as to certain mortgage characteristics and state-level foreclosure laws.

CGP finds a 27 percent foreclosure discount for single-family properties in Massachusetts over a significantly longer period of time, 1987–2007. In addition to looking at foreclosures, the authors identify properties in which the owners have died or declared bankruptcy. For these properties, they estimate a much lower discount: between 3 and 7 percent. CGP's authors find evidence that the death discount is due to lack of maintenance while the bankruptcy discount is due to liquidity concerns. They find evidence that the foreclosure discount is larger for houses in poorer neighborhoods with less valuable characteristics and interpret this as evidence that the discount is related to concerns about vandalism: Either the properties have been damaged while they are on the market or the cost of protection against potential vandalism is so high that the mortgage lenders are willing to accept significant discounts to sell them quickly. This difference in the estimated forced

sales discounts between foreclosures and deaths/bankruptcies is consistent with the investment externality effect that we discussed above. Since the non-foreclosure-related forced sales do not suffer from a separation between physical and legal ownership, they are less likely to experience the misaligned incentives that could impede property maintenance. This could explain the relatively small discount that CGP finds for non-foreclosure related forced sales compared to foreclosures.<sup>18</sup>

Second, the literature on subsidized investment, discussed in Section 4, shows that investment in property affects the value of nearby properties. Both Schwartz et al. (2006) and Rossi-Hansberg, Sarte, and Owens III (2010) examine programs that encouraged investment in individual properties and find important effects. Combining the findings of the foreclosure discount literature and the investment externalities literature, it would almost be surprising if previous studies had not uncovered foreclosure externalities.

What specific evidence do we find in favor of the investment externalities theory? First, there is the evidence on the condition of REO property. Again, it is hard to reconcile the demand or supply stories with the fact that the coefficient estimate associated with nearby below average REO comes in so far below the coefficient estimate associated with nearby REO in average and above average condition. Second, the investment externalities story can explain why the estimate is negative even for minor delinquencies and why it disappears after the new arms-length owner has had a chance to invest in the property.

---

<sup>18</sup>Clauret and Daneshvary (2009) also finds a foreclosure discount of slightly less than 10 percent in the Las Vegas metropolitan area between 2004 and 2007. In contrast, Harding, Rosenblatt, and Yao (2011) finds little evidence of a meaningful foreclosure discount, using both hedonic and repeat-sale methods.

## 6 Conclusion

In this paper, we document some new facts about foreclosure externalities. We show that houses trade at slightly lower prices when there are homes nearby with delinquent homeowners, when there are homes nearby owned by lenders, and even when there are homes nearby recently sold by lenders in arm's length transactions. We show that nearby houses trade at lower prices when the lender-owned property is in below-average condition and at higher prices when it is in above average condition. We discuss three possible explanations for the facts: supply effects, demand effects, and investment externalities, and argue that the third is the most plausible. But to be sure, the interactions of all three are so complex that no dataset and no model could likely completely rule out an explanation or precisely allocate the observed correlations to one of the three stories.

Perhaps the most important conclusion that one should take from this analysis is that the effects of foreclosure and distressed property in general on the prices of neighboring homes are fairly small. We estimate the effect of a property in SDQ and a property in REO on the price of a home within 0.10 mile to be approximately -0.5 to -1.0 percent, an amount that would most likely go unnoticed by the typical seller who does not have many distressed homeowners living nearby. As Table 4 shows, the vast majority of properties sell without any distressed property nearby, meaning that it is impossible to attribute more than a token amount of the collapse in prices in the 2006–2010 period to foreclosures. Our estimates of the negative externality are smaller than the estimates in much of the previous literature, and this likely has to do with differences in empirical strategies. Our identification of the externality is based on using within-CBG variation in price growth and nearby distressed properties, which we argue is more robust to reverse causality and simultaneity concerns than most previous studies in the literature.

Finally, the policy implications of even a small investment externality effect are important, especially in many of the areas that have been char-

acterized by large numbers of distressed properties throughout the recent foreclosure crisis. Our results suggest that the key to minimizing the costs of foreclosure is to minimize the time that properties spend in serious delinquency and in REO. On one hand, this implies putting pressure on lenders to sell properties out of REO quickly. On the other hand, and perhaps much less palatably, it implies minimizing the time a borrower spends in serious delinquency, which means accelerating the foreclosure process.

Put another way, our results suggest that delaying the foreclosure process exacts a cost on society as a whole that should be taken into account when making policy. As an example, Massachusetts passed a “right-to-cure” law in 2007, which forced lenders to give borrowers an additional 90 days to cure their mortgage before foreclosure proceedings could start. Gerardi, Lambie-Hanson, and Willen (2011) use a difference-in-differences approach to show that the law did not benefit borrowers in the sense that borrowers subject to the law were no more likely to cure or to renegotiate their loans than borrowers who were not. One might say that the law only failed to produce benefits, but our analysis suggests that it may also have imposed costs on homeowners who lived near borrowers who were able to take advantage of the law.

## References

- Anenberg, Eliot, and Edward Kung. 2012. "Estimates of the Size and Source of Price Declines due to Nearby Foreclosures." Working paper. [https://sites.google.com/site/elliotanenberg/research/Text\\_AK.pdf](https://sites.google.com/site/elliotanenberg/research/Text_AK.pdf).
- Campbell, John Y, Stefano Giglio, and Parag Pathak. 2011. "Forced Sales and House Prices." *American Economic Review* 101(5): 2108 – 2131.
- Clauret, T.M., and N. Daneshvary. 2009. "Estimating the House Foreclosure Discount Corrected for Spatial Price Interdependence and Endogeneity of Marketing Time." *Real Estate Economics* 37(1): 43–67.
- Frame, W.S. 2010. "Estimating the Effect of Mortgage Foreclosures on Nearby Property Values: A Critical Review of the Literature." *FRB Atlanta Economic Review*.
- Galster, G.C., P. Tatian, and R. Smith. 1999. "The Impact of Neighbors Who Use Section 8 Certificates on Property Values." *Housing Policy Debate* 10(4): 879–917.
- Gerardi, K., A.H. Shapiro, and P.S. Willen. 2009. "Decomposing the Foreclosure Crisis: House Price Depreciation Versus Bad Underwriting." *Working Paper*.
- Gerardi, Kristopher, Lauren Lambie-Hanson, and Paul S. Willen. 2011. "Do Borrower Rights Improve Borrower Outcomes? Evidence from the Foreclosure Process." Working Paper 17666. National Bureau of Economic Research.
- Gerardi, Kristopher, Adam Hale Shapiro, and Paul S. Willen. 2007. "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures." Working Paper 07-15. Boston, MA: Federal Reserve Bank of Boston.  
Available at <http://www.bos.frb.org/economic/wp/wp2007/wp0715.pdf>.
- Harding, J.P., E. Rosenblatt, and V. Yao. 2011. "The Foreclosure Discount: Myth or Reality?" Forthcoming *Journal of Urban Economics*.

- Harding, J.P., E. Rosenblatt, and V.W. Yao. 2009. "The Contagion Effect of Foreclosed Properties." *Journal of Urban Economics* 66(3): 164–178.
- Hartley, Daniel. 2011. "The Effect of Foreclosures on Nearby Housing Prices: Supply or Disamenity?" Working Paper 10-11R. Federal Reserve Bank of Cleveland.
- Immergluck, D., and G. Smith. 2006. "The External Costs of Foreclosure: The Impact of Single-family Mortgage Foreclosures on Property Values." *Housing Policy Debate* 17(1): 57–79.
- Lin, Z., E. Rosenblatt, and V.W. Yao. 2009. "Spillover Effects of Foreclosures on Neighborhood Property Values." *The Journal of Real Estate Finance and Economics* 38(4): 387–407.
- Linden, L., and J.E. Rockoff. 2008. "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws." *The American Economic Review* 98(3): 1103–1127.
- Pennington-Cross, A. 2006. "The Value of Foreclosed Property." *Journal of Real Estate Research* 28(2): 193–214.
- Rogers, W.H., and W. Winter. 2009. "The Impact of Foreclosures on Neighboring Housing Sales." *Journal of Real Estate Research* 31(4): 455–479.
- Rossi-Hansberg, E., P.D. Sarte, and R. Owens III. 2010. "Housing Externalities." *Journal of Political Economy* 118(3).
- Schuetz, J., V. Been, and I.G. Ellen. 2008. "Neighborhood Effects of Concentrated Mortgage Foreclosures." *Journal of Housing Economics* 17(4): 306–319.
- Schwartz, A.E., I.G. Ellen, I. Voicu, and M.H. Schill. 2006. "The External Effects of Place-based Subsidized Housing." *Regional Science and Urban Economics* 36(6): 679–707.

Table 2: Distribution of Repeat-Sale Observations By MSA and Year

**Panel A: By Metropolitan Statistical Area**

	# CBGs	Share of repeat-sale Transactions					Total
		2006	2007	2008	2009	2010	
Atlanta	1,020	26.8%	23.0%	17.7%	14.9%	17.6%	71,086
Boston	1,029	19.5	21.0	18.1	21.0	20.4	22,897
Chicago	1,139	24.2	19.7	16.9	20.3	19.0	27,853
Detroit	1,060	18.3	20.1	20.9	20.7	20.0	23,300
LV	438	25.6	15.1	17.2	20.5	21.6	71,189
LA	2,228	24.3	16.3	16.1	22.1	21.2	84,552
Miami	825	33.1	19.6	12.8	16.3	18.1	67,240
New York	2,023	21.5	22.3	19.7	20.4	16.2	45,305
Orlando	383	29.3	17.7	13.2	17.9	22.0	29,906
Philadelphia	1,485	24.7	23.4	18.0	18.3	15.8	51,038
Phoenix	1,069	26.1	18.3	17.1	18.8	19.7	154,243
Riverside	885	27.8	14.9	16.2	19.7	21.4	94,435
Seattle	836	28.5	24.6	16.1	15.7	15.1	35,893
Tampa	745	30.0	18.7	14.5	18.1	18.6	54,743
DC	1,767	26.2	19.2	16.7	19.4	18.5	124,833
Total	16,932	252,478	181,551	159,176	181,391	183,917	958,513

**Panel B: By Year**

Year of Purchase	Year of Sale					Total	%
	2006	2007	2008	2009	2010		
2001	30,714	19,135	12,243	11,846	11,488	85,426	9.0%
2002	37,600	23,699	14,942	14,673	14,279	105,193	11.0%
2003	49,441	30,042	19,476	19,220	18,486	136,665	14.3%
2004	58,144	37,750	25,166	25,203	24,454	170,717	17.9%
2005	56,544	34,501	34,412	36,221	34,171	195,849	20.6%
2006	19,966	24,237	28,386	32,574	29,500	134,663	14.1%
2007	.	12,112	14,690	14,996	17,030	58,828	6.2%
2008	.	.	9,741	13,069	7,508	30,318	3.2%
2009	.	.	.	12,369	10,696	23,065	2.4%
2010	.	.	.	.	11,704	11,704	1.2%
Total	252,409	181,476	159,056	180,171	179,316	952,428	100.0%
%	26.5%	19.1%	16.7%	18.9%	18.8%	100.0%	

Notes: This table reports the distribution of repeat-sale pairs of non-distressed transactions across the 15 largest metropolitan statistical areas (MSAs) in the United States. The data are drawn from public records purchased from a national data aggregator. The data include all repeat-sale pairs in which the first sale occurred between 2001 and 2010 and the second sale between 2006 and 2010.

Table 3: Summary Statistics of Property Characteristics for Repeat-Sales Observations

	Mean	Std	Min	p25	p50	p75	Max
Price (\$)	331,217	252,565	4,500	180,900	270,400	400,000	9,775,000
Prior (\$)	349,559	279,454	4,500	187,000	285,000	425,000	9,500,000
Appreciation (%)	2.9	43.3	-204.0	-19.8	5.7	32.1	187.2
Holding Period (qtrs)	13.9	8.4	1	8	13	20	39
<i>Property Characteristics</i>							
Size (square ft.)	1,973	849	500	1,372	1,770	2,371	32,222
Lot (square ft.)	12,424	23,014	500	5,200	7,405	10,890	435,600
Age (years)	23.8	21.9	0	7	16	36	150

Notes: This table displays summary statistics of the hedonic variables that are included in the various empirical models discussed in the text. The mean, standard deviation, minimum, 25th percentile (p25), 50th percentile (p50), 75th percentile (p75), and maximum of each variable are displayed.



Table 4: Summary Statistics of Nearby Distressed Properties for Repeat-Sales Observations

<b>A: % of sales with # distressed property within 330 feet</b>								
		0	1	2	3	4	≥5	
# SDQs		65.6	19.6	7.7	3.5	1.7	2.0	
	Long SDQs	87.6	8.4	2.4	0.9	0.4	0.4	
	Short SDQs	70.7	19.3	6.4	2.2	0.8	0.6	
	Minor DQs	30.2	18.8	13.9	10.4	7.6	19.0	
REO Inventory		81.3	11.9	3.9	1.6	0.7	0.7	
REO Sales in last year		80.4	10.3	3.9	2.0	1.2	2.1	
REO Sales 1-2 years ago		86.5	8.2	2.5	1.2	0.6	1.0	
Total Distressed		66.7	14.9	6.5	3.7	2.3	6.0	
<b>B: % of sales with difference in # distressed property within 330 feet</b>								
		< 0	0	1	2	3	4	≥ 5
Long SDQs		1.5	84.5	9.1	2.9	1.1	0.5	0.4
REO Inventory		4.8	77.8	9.3	3.5	1.8	1.0	1.8
REO Sales in last year		5.2	83.0	7.2	2.2	1.0	0.6	0.9
REO Sales 1-2 years ago		0.2	89.3	7.1	2.0	0.8	0.3	0.3
<b>C: % of Sales with &gt; 0...</b>								
		2006	2007	2008	2009	2010	Total	
SDQs		22.9	22.8	29.9	45.0	55.5	34.5	
	Long SDQs	1.7	2.1	6.5	23.7	31.2	12.4	
	Short SDQs	22.0	21.7	26.3	33.7	44.9	29.3	
	Minor DQs	68.6	68.6	68.8	70.8	72.2	69.8	
REO Inventory		7.4	10.2	23.9	27.1	29.7	18.7	
REO Sales in last year		3.8	4.1	18.1	38.4	39.5	19.6	
REO Sales 1-2 years ago		6.1	3.8	4.5	17.5	37.1	13.5	
Total Distressed		11.7	14.5	35.9	54.3	58.8	33.4	
<b>D: Correlations</b>								
	SDQ inventory (< 0.1 miles)	REO inventory (< 0.1 miles)	REO sales in last year (< 0.1 miles)	REO sales 1-2 years ago (< 0.1 miles)	SDQ inventory (0.1-0.25 miles)	REO inventory (0.1-0.25 miles)	REO sales in last year (0.1-0.25 miles)	REO sales 1-2 years ago (0.1-0.25 miles)
SDQ inventory (< 0.1 miles)	1							
REO inventory (< 0.1 miles)	0.31	1						
REO sales in last year (< 0.1 miles)	0.44	0.47	1					
REO sales 1-2 years ago (< 0.1 miles)	0.33	0.29	0.45	1				
SDQ inventory (0.1 - 0.25 miles)	0.61	0.30	0.45	0.35	1			
REO inventory (0.1 - 0.25 miles)	0.31	0.65	0.49	0.31	0.40	1		
REO sales in last year (0.1 - 0.25 miles)	0.44	0.46	0.73	0.45	0.55	0.59	1	
REO sales 1-2 years ago (0.1 - 0.25 miles)	0.34	0.29	0.45	0.67	0.43	0.39	0.56	1

Notes: This table displays summary statistics on the number of nearby distressed properties to the non-distressed repeat-sales observations in the sample. A “short” SDQ is defined as a property in which the borrower has been in serious delinquency for less than one year, a “long” SDQ is defined as a serious delinquency that has lasted for more than one year, and a “minor delinquency” is defined as a property in which the borrower is only one or two payments behind. Total Distressed refers to long SDQs, REO inventory, and both REO sales variables.

Table 5: Baseline Specification. Dependent variable is price growth between purchase and sale.

Panel A: Baseline Specification				Panel B: Different Years				
	(1)	(2)	(3)	2006	2007	2008	2009	2010
$\Delta$ Long SDQ (< 0.1 miles)	-0.012 (24.6)	-0.011 (22.3)	-0.012 (23.6)	-0.052 (14.3)	-0.032 (10.2)	-0.018 (14.3)	-0.012 (24.6)	-0.015 (33.3)
$\Delta$ Foreclosure Inventory (< 0.1 miles)	-0.009 (21.5)	-0.008 (19.5)	-0.008 (19.8)	-0.019 (18.3)	-0.004 (5.1)	-0.007 (14.8)	-0.009 (21.5)	-0.010 (22.3)
$\Delta$ REO sold in last year (< 0.1 miles)	-0.006 (68.9)	-0.005 (66.6)	-0.005 (71.2)	-0.001 (3.5)	-0.003 (4.9)	-0.006 (15.0)	-0.006 (24.6)	-0.006 (23.6)
$\Delta$ REO sold 1-2 years ago (< 0.1 miles)	0.001 (1.6)	0.001 (2.2)	0.001 (2.7)	-0.001 (1.9)	-0.002 (3.3)	0.000 (0.5)	0.001 (1.6)	-0.001 (3.1)
$\Delta$ Long SDQ (0.1 – 0.25 miles)	-0.008 (43.9)	-0.008 (44.0)	-0.008 (44.8)	-0.045 (22.4)	-0.025 (15.6)	-0.016 (28.5)	-0.008 (43.9)	-0.008 (54.4)
$\Delta$ Foreclosure Inventory (0.1 – 0.25 miles)	-0.007 (40.2)	-0.006 (39.8)	-0.007 (40.1)	-0.007 (17.4)	0.002 (5.9)	-0.003 (16.6)	-0.007 (40.2)	-0.008 (45.9)
$\Delta$ REO sold in last year (0.1 – 0.25 miles)	-0.002 (23.3)	-0.002 (24.9)	-0.002 (26.1)	-0.001 (7.6)	0.000 (1.9)	-0.003 (18.1)	-0.002 (23.3)	-0.002 (22.1)
$\Delta$ REO sold 1-2 years ago (0.1 – 0.25 miles)	0.004 (28.7)	0.004 (28.2)	0.004 (29.0)	0.000 (1.7)	-0.001 (4.8)	0.003 (14.6)	0.004 (28.7)	0.002 (18.2)
Property Age	0.001 (23.0)	0.001 (24.8)	-	$R^2$				
Property Age <sup>2</sup>	0.000 (18.0)	0.000 (19.3)	-	0.39	0.44	0.69	0.84	0.80
Log(Lot size)	-0.004 (4.8)	-0.015 (23.2)	-	Panel C: Different Geographic Controls				
Log(Living area)	0.068 (42.9)	0.064 (40.5)	-	None	MSA	County	Tract	CBG
# of properties <330 ft	0.837 (26.4)	-	-	-0.023 (36.5)	-0.016 (29.1)	-0.017 (30.6)	-0.012 (24.5)	-0.012 (24.6)
$R^2$	0.84	0.83	0.82	-0.019 (35.7)	-0.013 (27.1)	-0.012 (25.5)	-0.010 (22.2)	-0.009 (21.5)
				-0.011 (37.6)	-0.009 (33.7)	-0.007 (27.6)	-0.006 (24.1)	-0.006 (24.6)
				-0.001 (1.4)	0.000 (0.0)	0.002 (4.0)	0.001 (1.8)	0.001 (1.6)
				-0.010 (-46.5)	-0.008 (43.6)	-0.010 (51.6)	-0.008 (43.9)	-0.008 (43.9)
				-0.013 (64.4)	-0.008 (45.7)	-0.009 (48.3)	-0.007 (41.9)	-0.007 (40.2)
				-0.005 (40.0)	-0.004 (36.0)	-0.003 (32.4)	-0.002 (23.2)	-0.002 (23.3)
				0.004 (21.9)	0.005 (28.5)	0.005 (33.1)	0.004 (29.0)	0.004 (28.7)
				$R^2$	0.33	0.61	0.76	0.83
					0.61	0.76	0.83	0.84

Notes: This table displays results from the estimation of equation (2). In all regressions reported, the independent variable is price appreciation over the repeat-sale interval. Panel A includes a full set of  $CBG \times \text{year-of-purchase}$  fixed effects. Panel B includes the hedonic controls listed in Panel A as well as  $CBG \times \text{year-of-purchase}$  fixed effects. In panels A and C the second sale year in the repeat-sale pair is 2009. The numbers in parentheses are  $t$  statistics.

Table 6: Effect of Property Condition and Vacancy Status based on 2010 Sample.

<i>Diff. in SDQs by Vacancy Status</i>	
Occupied	-0.006 (7.4)
Vacant	-0.011 (13.9)
Missing Data	-0.008 (9.9)
<i>Diff. in REO in Inv by Property Condition</i>	
Below average	-0.015 (5.0)
Average	-0.001 (0.8)
Above average	0.023 (6.00)
Missing Data	-0.005 (10.2)
Diff. in REO sold in past year	-0.016 (38.3)
Diff. in REO sold 1-2 year ago	-0.007 (25.4)
$R^2$	0.77

Note: All regressions include the specification in column (1) of Panel A in Table 5, including controls for the number of distressed properties between 0.10 and 0.25 mile (not shown), and a full set of CBG $\times$ year-of-purchase dummies. The top panel in the table displays results for close (within 0.10 mile) SDQ inventory by vacancy status, while the bottom panel displays results for close REO inventory in different states of upkeep. Numbers in parentheses are  $t$  statistics.

Table 7: Effect of REO/SDQs by Distance based on 2009 Sample.

	Long SDQs	Difference in number of: REO inventory	REO sold in:	
			Last year	1-2 years ago
0.0 - 0.10 miles	-0.007 (14.3)	-0.007 (17.4)	-0.006 (24.6)	-0.001 (2.4)
0.10 - 0.25 miles	-0.003 (9.5)	-0.002 (15.6)	-0.001 (9.2)	0.002 (13.8)
0.25 - 0.50 miles	-0.001 (21.9)	-0.002 (12.3)	-0.000 (7.0)	0.001 (13.2)
0.50 - 1.0 miles	-0.001 (30.4)	-0.001 (15.3)	0.000 (4.1)	0.001 (16.2)
$R^2$		0.87		

Note: The regression is specified as in column (1) of Panel A in Table 5 and a full set of CBG×year-of-purchase dummies. Numbers in the parenthesis are  $t$  statistics.

Table 8: Alternative Specifications.

	(1)	(2)	Baseline		(5)	(6)	RHS in levels	Hedonic
			(3)	(4)			(7)	(8)
Minor DQs			-0.005 (19.0)					
Short SDQs		-0.011 (18.6)	-0.007 (12.3)					
Long SDQs	-0.013 (25.9)	-0.011 (22.7)	-0.012 (24.4)				-0.009 (18.4)	-0.013 (20.9)
All SDQs					-0.011 (31.1)			
All SDQs with infill						-0.009 (31.3)		
REO inventory	-0.009 (22.1)	-0.009 (21.6)	-0.010 (23.1)	-0.009 (22.1)	-0.009 (21.6)	-0.007 (17.9)	-0.005 (12.4)	-0.008 (15.7)
REO sold in past year	-0.006 (25.4)	-0.006 (24.6)	-0.007 (27.1)	-0.007 (28.8)	-0.006 (24.6)	-0.005 (22.2)	-0.006 (24.0)	-0.008 (25.3)
REO sold 1-2 year ago	0.000 (1.2)	0.001 (1.9)	0.000 (1.4)	0.000 (1.0)	0.001 (1.8)	0.000 (1.0)	-0.003 (5.0)	-0.004 (6.8)
$R^2$	0.73	0.73	0.73	0.72	0.73	0.73	0.71	0.84

Note: All regressions include the specification in column (1) of Panel A in Table 5, including controls for the number of distressed properties between 0.10 and 0.25 mile (not shown) and a full set of CBG $\times$ year-of-purchase dummies. All reported coefficients are significant at the 1% significance level. Numbers in parentheses are  $t$  statistics.

Figure 1: Census Tracts and Blocks in Cambridge, MA. 2010 Census.

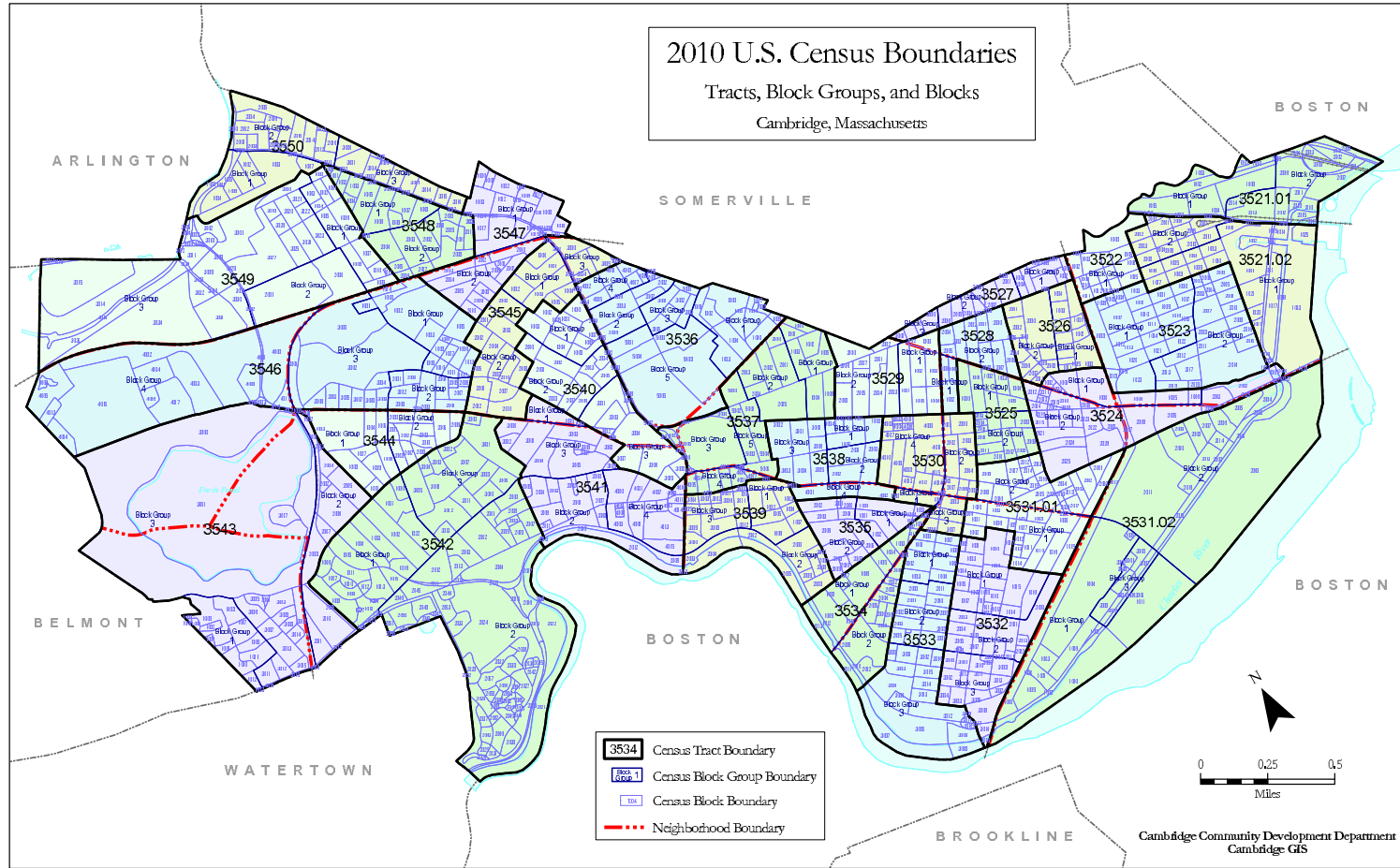
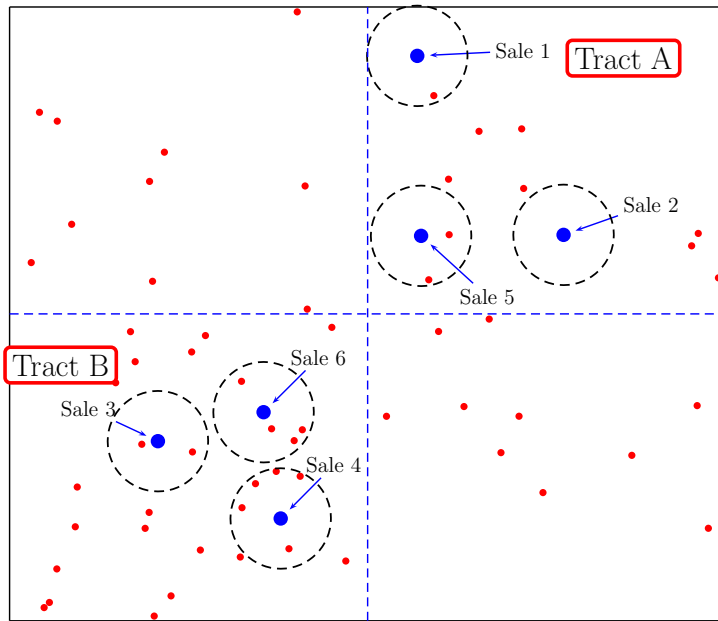


Figure 2: Importance of Geographic Controls

### Map of Foreclosures and Sales



### Map of Foreclosures and Sales *without* Census Tract Controls

