Forced Sales and House Prices

John Y. Campbell, Stefano Giglio, and Parag Pathak[1]

First draft: September 2008
This version: December 2009

[1]Campbell: Department of Economics, Littauer Center, Harvard University, Cambridge MA 02138, USA, and NBER. Email john_campbell@harvard.edu. Giglio: Department of Economics, Littauer Center, Harvard University, Cambridge MA 02138, USA, and visiting scholar, Federal Reserve Bank of Boston. Email sgiglio@fas.harvard.edu. Pathak: Department of Economics, 50 Memorial Drive, E52 MIT, Cambridge, MA 02142. Email: ppathak@mit.edu. We are grateful to Tuomo Vuolteenaho and Paul Willen for early conversations which stimulated our thinking on this topic, to the Federal Reserve Bank of Boston and the Lincoln Institute of Land Policy for assistance in obtaining data, to the Real Estate Academic Initiative at Harvard for financial support, and to Ed Glaeser, Bob Hall, David Laibson, Jeff Pontiff, Tomasz Piskorski, Stuart Rosenthal, James Vickery, Susan Woodward, and seminar participants at the Federal Reserve Bank of Boston, the Federal Reserve Board, the University of California at Berkeley, Stanford University, Columbia University, MIT, the NBER Summer Institute, Brown University, the London School of Economics, and London Business School for comments on an earlier draft.

Abstract

This paper uses data on house transactions in the state of Massachusetts over the last 20 years to show that houses sold after foreclosure, or close in time to the death or bankruptcy of at least one seller, are sold at lower prices than other houses. Foreclosure discounts are particularly large on average at 27% of the value of a house. The pattern of death-related discounts suggests that they may result from poor home maintenance by older sellers, while foreclosure discounts appear to be related to the threat of vandalism in low-priced neighborhoods. After aggregating to the zipcode level and controlling for regional price trends, the prices of forced sales are mean-reverting, while the prices of unforced sales are close to a random walk. At the zipcode level, this suggests that unforced sales take place at approximately efficient prices, while forced-sales prices reflect time-varying illiquidity in neighborhood housing markets. At a more local level, however, we find that foreclosures that take place within a quarter of a mile, and particularly within a tenth of a mile, of a house lower the price at which it is sold. Our preferred estimate of this effect is that a foreclosure at a distance of 0.05 miles lowers the price of a house by about 1%.

1 Introduction

The market for housing differs in several important ways from the textbook model of a liquid asset market with exogenous fundamentals. This implies that the price at which a house is sold can be influenced not only by general supply and demand conditions, but also by idiosyncratic factors including the urgency of the sale and the effects of the ownership transfer on the physical quality of the house.

First, houses are productive only when people are living in them. Owning an empty house is equivalent to throwing away the dividend on a financial asset. Second, houses are fragile assets that need maintenance, and are vulnerable to vandalism. Unoccupied houses are particularly vulnerable and expensive to protect. Third, short-term rental contracts involve high transactions costs, resulting from the moving costs of renters and the need of homeowners to protect their property against damage. Fourth, houses are expensive, indivisible, and heterogeneous assets. Each house has certain unique characteristics which are likely to appeal to certain potential buyers and not to others, so selling a house requires matching it with an appropriate buyer. Because of the high costs of intermediation in housing, this task is normally undertaken by a real estate broker rather than a dealer. Fifth, most homeowners must finance their purchases using mortgages, collateralized debt contracts that transfer home ownership to the mortgage lender through a foreclosure process if the homeowner defaults.

The expansion of mortgage credit earlier this decade and the recent decline in house prices have led to an unprecedented increase in foreclosures since 2006. Foreclosures transfer houses to financial institutions who must maintain and protect them until they can be sold. Foreclosed houses are likely to sell at low prices, both because they may have been physically damaged during the foreclosure process, and because financial institutions have an incentive to sell them quickly. In a liquid market, an asset can be sold rapidly with a minimal impact on its price, but the characteristics of housing discussed above make the market for residential real estate a classic example of an illiquid market, in which urgent sales lower prices.[2]

There is widespread concern that foreclosures may also lower the prices of nearby houses, either through direct physical effects on neighborhoods or by creating an imbalance of demand and supply in an illiquid neighborhood housing market. If such spillover effects on prices are important, they might stimulate further foreclosures because homeowners are more likely to default when their houses are worth less than the face value of their mortgages. See for example the motivation for the Obama Administration's *Making Home Affordable* plan, as described on the US Treasury website: "In the absence of decisive action, we risk an intensifying spiral in which lenders foreclose, pushing area home

[2] Mayer (1995) presents a theoretical model of this effect, assuming that an urgent sale is implemented using an auction.

prices still lower, reducing the value of household savings, and making it harder for all families to refinance. In some studies, foreclosure on a home has been found to reduce the prices of nearby homes by as much as 9%." (US Treasury 2009.)

In this paper we seek to understand the illiquidity of the housing market, and specifically the effects of foreclosures on the prices of foreclosed houses and other houses in the same neighborhood. We use a comprehensive dataset on individual house transactions in Massachusetts over the period from 1987 through the first quarter of 2009. Importantly, Massachusetts experienced a significant decline in house prices and wave of foreclosures during the early 1990s, which gives us a historical precedent that can be used to shed light on the current condition of the housing market.

We study several categories of sales which plausibly are more urgent than normal. We first link data on house transactions in the state of Massachusetts, over the period 1987 to March 2009, to information on deaths and bankruptcies of individuals. By matching names and addresses across datasets, we are able to identify transactions as forced sales if they occur close in time to the death or bankruptcy of at least one seller. We use hedonic regressions with neighborhood fixed effects, standard in the real estate literature, to control for heterogeneity in the characteristics of houses. We find that forced sales take place at price discounts of about 3-7%, and these discounts increase when a house has one seller rather than two.

One concern about this finding is that it might reflect unobserved effects of death or bankruptcy on the quality of a house, in particular deferred maintenance by homeowners with health or financial problems. In order to explore this issue, we examine how discounts vary with the timing of sales in relation to the seller's death or bankruptcy, we separate the deaths of younger and older sellers, we distinguish housing types, and we relate discounts to the various components of a property's value. We find that death-related discounts are not closely related to the timing of a sale in relation to death, are larger for older sellers, smaller for condominiums, and larger for houses whose structures account for a larger fraction of their value. This evidence suggests that death-related discounts reflect poor maintenance of houses by older sellers, while bankruptcy-related discounts appear more closely related to the urgency of sale immediately after bankruptcy.

Our main interest is in foreclosures. We find large foreclosure discounts, about 27% on average. These discounts are not highly sensitive to the type of housing, but they are larger for houses with low-priced characteristics in low-priced neighborhoods. This suggests that the foreclosure discount may be related to vandalism, through two possible channels. First, foreclosed houses may have been damaged before they are sold. Second, mortgage lenders must protect foreclosed houses while they are vacant; the threat of vandalism may be greater in bad neighborhoods, and costs of protection likely account for a larger fraction of the value of a low-priced house. The costs of protection induce

mortage lenders to sell foreclosed houses urgently, leading to discounts in illiquid housing markets.

The incidence of foreclosure sales is highly variable over time and space, but in some areas at some times foreclosures account for a large fraction of total sales. This allows us to study the relations between forced sales prices and the subsequent transactions prices of other houses in the same neighborhood.

We contrast two extreme views of the relation between forced and unforced sales prices for houses. The first view is that unforced transactions take place at efficient prices, which evolve following a random walk, while forced sales take place at lower prices. If the housing market were a dealer market with a bid-ask spread, we could think of unforced transactions as revealing the efficient price at the midpoint of the spread, while forced transactions reveal the lower bid price. If the bid-ask spread is variable over time, then large discounts of forced from unforced sales prices should predict increases in forced sales prices, but should have no implications for future prices of unforced transactions. That is, bid-ask bounce (Roll 1984) affects the prices of forced sales but not those of unforced sales.

The opposite extreme view is that forced sales convey information about the future prices of unforced transactions. There are several reasons why this might be the case. First, forced sales may perform the function of price discovery, revealing the prices at which buyers are willing to enter the market. Particularly in down markets, homeowners without urgent motives to sell may set unrealistically high prices, perhaps because their expectations lag the market or because they use their purchase price as a reference price (Genesove and Mayer 2001). In this situation, unforced transactions may take place only when particularly enthusiastic buyers appear. If the housing market had a bid-ask spread, we could think of forced transactions as revealing the efficient price at the midpoint of the spread, while unforced transactions reveal the higher ask price. If the bid-ask spread varies over time, a large discount of forced from unforced prices would predict declines in unforced sales prices.

There could also be causal effects of forced sales on the general level of house prices. Forced sales could absorb demand, reducing the prices of those houses that come to market later. Forced sales could affect the reference prices that buyers and sellers use as "comparables" when they negotiate prices. In the case of foreclosures, there is widespread concern that there may be direct negative effects of foreclosures on neighborhoods. Foreclosures typically involve periods during which houses stand empty, reducing the visual appeal and social cohesion of the neighborhood and encouraging crime (Apgar, Duda, and Gorey 2005, Immergluck and Smith 2005, 2006).

Despite the plausibility of these concerns, we find that at the zipcode level, the prices of forced sales have relatively little predictive power for the prices of other transactions in the housing market. The discount between urgent sales prices and other sales prices is stationary, so when it widens, it

normally narrows again. But this primarily occurs through an increase in the prices of forced sales, not through a decrease in the prices at which other transactions occur.

In order to detect spillover effects from forced sales to unforced sales, we look at foreclosures that take place within a quarter of a mile, and within a tenth of a mile, of each transaction in our dataset. At this highly local level, we do see evidence that foreclosures lower house prices, and the effect is economically significant during foreclosure waves. The extremely localized nature of these spillover effects is consistent with results reported by Harding, Rosenblatt, and Yao (2008) for foreclosures, and by Rossi-Hansberg, Sarte, and Owens (2008) for urban revitalization expenditures. The spillover effects of foreclosures are persistent and, like the discounts on foreclosed houses, they are larger in low-priced neighborhoods. Both results suggest that spillovers may reflect physical damage to neighborhoods.

The forced sale discounts we report in this paper are consistent with earlier findings of illiquidity in the housing market. There is evidence that certain seller characteristics influence selling price and time on the market in the same direction, as would be expected if an urgent desire to sell lowers the price that a house fetches. Genesove and Mayer (1997) show that homeowners with larger mortgages relative to their home values set higher asking prices, realize higher prices if they sell, but keep their homes on the market longer than homeowners with smaller mortgages. More precisely, they find that a house with a loan-to-value ratio of 100% sells for 4% more but stays on the market 15% longer than a house with a loan-to-value ratio of 80%. Levitt and Syverson (2008) show that realtors selling their own houses get higher prices and keep their homes on the market longer than their clients do. The price differential is about 4%, and the time on the market differential is about 10%, numbers which are roughly comparable to those reported by Genesove and Mayer. Mayer (1998) studies real estate auctions, which in the United States have been used primarily as a rapid sales mechanism by developers and banks, and finds discounts of up to 9% in Los Angeles during a real estate boom, and between 9% and 21% in Dallas during a real estate bust.

A related literature in corporate finance argues that assets with limited alternative uses appeal to relatively few buyers and are correspondingly less valuable when they must be urgently sold. This affects the debt contracts that can be used to finance such assets (Shleifer and Vishny 1992). Benmelech, Garmaise, and Moskowitz (2005) apply this insight to commercial real estate.

The organization of the paper is as follows. Section 2 describes our data and the procedures we have used to clean it. Section 3 presents our hedonic regression methodology and uses it to estimate the discounts of forced sales from unforced sales. This section also uses cross-sectional variation in discounts to distinguish alternative interpretations. Section 4 studies the ability of forced and unforced sales prices to predict future changes in house prices within the same zip codes, and more

local spillover effects from foreclosures to house prices in the immediate neighborhood. Section 5 concludes.

2 House Price and Forced Sale Data

2.1 House prices

We begin with a dataset on changes in ownership of residential real estate, provided to us by the Warren Group. The data cover the period 1987 to March 2009, and the entire state of Massachusetts. The online appendix to this paper (Campbell, Giglio, and Pathak 2009) shows the number of transactions by zip code to illustrate the geographical coverage of the data.

The Warren Group data record basic characteristics of the houses involved in each transaction. In almost all cases, the characteristics are measured as of August 2007; about 78,000 houses were added to the dataset after this date and have characteristics measured later. Unfortunately, we do not have a dynamic dataset tracking changes in house characteristics over time.[3]

The Warren Group data also record the sales price of each house and the names of buyers and sellers. We have carefully cleaned the data to remove transactions that appear to be intra-family transfers of ownership rather than arms-length transactions, and duplicate transactions that reflect intermediation or corrections of public records. The online appendix describes our data cleaning procedures in detail.

We remove outliers from the Warren Group data in several steps. We exclude transactions in properties that cannot be classified as either single family, multifamily, or condominiums, and transactions that take place at extreme prices, below the 1st or above the 99th percentile of the distribution of raw prices. Where the dataset reports impossible property characteristics (for example, zero rooms), we treat these characteristics as missing. Finally, we winsorize reported square footage at the 1st and 99th percentiles and reported numbers of rooms at the 99th percentile. The resulting dataset has 1,831,393 transactions.

The median house, across all transactions in all years, has 1,535 square feet of living area on a 9,452 square foot lot; it is 38 years old with 6 rooms, 3 bedrooms, and 2.0 bathrooms, and sells for a nominal price of $180,000. The means of these characteristics are slightly higher than the medians, indicating right skewness of the distribution, for all these characteristics. Full details on both house and census tract characteristics are presented in appendix Table A.1.

[3] One might be concerned that inaccurately measured housing characteristics early in our sample period could affect our results. However, in Tables A.6 and A.13 of the online appendix we find very similar results throughout our sample period.

2.2 Forced sales

In order to identify forced sales, we obtain data on deaths and bankruptcy filings from the Death Master File of the Social Security Administration and Lexis/Nexis, respectively. These data give us names, addresses, and dates which can be matched to the names and addresses of house sellers in the Warren Group data. Many houses have two joint sellers, and we classify the sale as forced if we can match the name of at least one of these sellers to a death or bankruptcy filing within three years of the house sale. The Death Master File also gives us the ages of sellers, information that is not available elsewhere in our dataset. Although our bankruptcy data include some corporate bankruptcy filings, only personal bankruptcies end up matched to house sales.

The algorithm we use for name matching is described in detail in the online appendix. We match based on last name, first name, and zip code. We then use sensible priority rules, based on match quality, middle initials, and event dates, to eliminate multiple matches.

We also identify forced sales related to foreclosures. Foreclosure proceedings typically begin after homeowners miss about three payments and are unable to negotiate a solution with their lenders. During this period, homeowners may be able to sell their property prior to actual foreclosure, but our data do not allow us to identify these cases. The Warren Group data report transfers of ownership that take place through foreclosure by demarcating the source of the transaction deed as foreclosure-related.

Massachusetts has both judicial and non-judicial foreclosures. A judicial foreclosure is processed through the courts, beginning with lender filing and recording a notice which includes the amount of outstanding debt and reasons for foreclosure. Non-judicial foreclosures, in contrast, are processed without court intervention, and the foreclosure requirements are established by state statutes. In either case, with assistance from the local sheriff's office, the first attempt at selling the property is via an auction. The trustee or attorney handling the foreclosure sets the opening bid and this is usually advertised in the foreclosure notice. The typical opening bid is the balance of the mortgage plus penalties, unpaid interest, attorney fees, and other costs that the lender has incurred during the process. In Massachusetts, the deposit to participate in the auction is usually $5,000 and homeowners are not obligated to allow bidders to investigate inside the property.[4]

Since Massachusetts does not have a redemption period where a homeowner retains the right to buy back the property by paying the full amount of the loan along with taxes, interest, and penalties, the transfer of ownership becomes complete at a closing following the foreclosure auction. The previous owners, if still present, are automatically converted to tenants, and the new owner must

[4] According to Massachusetts law, if there are two mortgages, the first of which forces the foreclosure, and there is no money left after the sale of the house to pay the second mortgage, the holder of the second mortgage still has a claim against the borrower, but no further claim against the house. However, in the relatively unusual case where a second lender forces foreclosure, the property is sold with a lien from the first mortgage.

follow Massachusetts legal procedures for eviction.[5]

Foreclosure auctions may be successful or unsuccessful. In a successful auction, the property is sold to the highest bidder at a price equal to or exceeding the opening bid. Successful auctions represent 18% of our cases. We identify these as cases where the acquirer is an individual or realty trust, or takes out a mortgage to finance the purchase.

In an unsuccessful auction, nobody bids higher than the opening bid and control is handed over to the lender. In this case, the lender is responsible for the sale of the property, and usually transfers the property to its real estate owned (REO) department, which prepares it for sale typically on the open market. Occasionally, REOs negotiate sales directly with investors rather than place the property on the market, and can even offer purchasers packages of properties. For these 82% of cases in our dataset, we treat the subsequent sale of the property by the mortgage lender as an urgent or forced sale.

In cases where a sale is both foreclosure-related and linked to a death or bankruptcy, we retain the foreclosure classification. If a sale is linked to both a death and a bankruptcy, we use priority rules, based on match quality and event dates, to classify it as either death-related or bankruptcy-related.

The top panel of Table 1 reports the frequency of each type of forced sale for each year in our data set. The first column of the table shows the total number of housing transactions in the Warren Group data in each year. We have just over 22 years of data and over 1.8 million transactions, for an average of just over 82,000 transactions per year. Of these, 6.1% are forced transactions: 3.5% related to foreclosures, 1.8% related to deaths, and 0.8% related to bankruptcies. The fraction of forced sales is highly variable over time. At the beginning and end of the sample, this is partially due to the matching process: we do not match deaths which happened before the start of our data or bankruptcies which occurred more than three years before the start date of our bankruptcy data in 1993. At the very end of the sample this is due to the fact that we cannot match sales to future deaths or bankruptcies. More generally, it reflects a gradual increase in death-related sales over time, and an upward shift in the incidence of bankruptcy in the late 1990s and early 2000s before bankruptcy reform increased the cost of personal bankruptcy in 2005.[6] However the most important time-variation is driven by two waves of foreclosures during the housing downturns of the early 1990s and 2007-09. The incidence of foreclosure-related forced sales was negligible in 1987, rose to 9.7% in 1993, then receded to under 1% in the mid-2000's before rising again to reach a record level of 25.7% in the first quarter of 2009.

[5] This can run anywhere from 6 weeks to 6 months, with the average about 10 weeks (http://www.lawlib.state.ma.us/foreclosure.html, "Foreclosure FAQ").

[6] Morgan, Iverson, and Botsch (2008) suggest that the bankruptcy reform of 2005 contributed to the subsequent increase in subprime mortgage defaults by making it harder for borrowers to achieve relief from unsecured debt obligations.

The bottom panel of Table 1 categorizes forced sales according to the date of the death, bankruptcy, or foreclosure in relation to the house sale. In the case of death, we find that house sales within one year of the death of a seller are more common than house sales two or three years before or after the death of a seller; however sales are almost equally common the year before a seller's death and the year after. In the case of bankruptcy, we find that house sales are relatively rare during the three years before a bankruptcy filing, but the sales incidence spikes up the year after the filing and then gradually declines. For instance, 30.8% of bankruptcy related sales take place the year after the bankruptcy filing, while only 9.5% take place the year before. The scarcity of sales before bankruptcy presumably reflects the fact that bankruptcy filing protects all but the most expensive primary residences from creditors through the homestead exemption (White 2008). Foreclosure-related sales cannot occur before the underlying foreclosure, and tend to take place rapidly thereafter. Of the 3.5% of foreclosure-related sales in our overall dataset, 85.9% occur within one year, 9.1% in the second year, 1.6% in the third year, and the remainder with a longer lag.

In the complete dataset, 65% of transactions are in single family houses, 11% in multi family houses, and 24% in condominiums. Among forced sales, however, multi family houses are more common (20%) and condominiums are less common (17%). The paper reports results both for the entire dataset, and for separate subsamples for each housing type.

The city of Boston accounts for 8% of all sales and almost 10% of forced sales. Boston's modestly greater share of forced sales is entirely caused by a higher incidence of foreclosures in Boston (13% of foreclosures are in the city). Death- and bankruptcy-related sales are actually less common in Boston than elsewhere. Figure 1 provides a richer picture of the geographic distribution of forced sales, plotting by zip code the share of forced sales in total sales.

When we compare the distribution of house characteristics for forced sales, we find that the median forced sales price takes place at $123,000, which is only two thirds of the median sales price in our overall dataset. This is true despite the fact that the median forced sale is of a similarly sized house on a lot 79% of the size of the median sale.

At first sight, the lower median price for forced sales suggests that these transactions take place at a large price discount. However, one cannot reach this conclusion based on this simple comparison. The incidence of forced sales was much greater in the early 1990s, when the overall level of prices was depressed; and forced sales are more likely to take place in low-income minority neighborhoods, where prices are likely to be lower for any given size of house.[7] The next step in our analysis is to control

[7] Table A.3 in the online appendix presents a comparison of house and neighborhood characteristics for forced sales relative to our overall dataset. We also estimated models where house characteristics are functions of four forced indicators – young death, old death, bankruptcy, and foreclosure– and census tract-year fixed effects. The estimates are presented in appendix Table A.4. The regression estimates indicate that forced sales tend to have between 0.10 and 0.19 more rooms

for these effects by using a hedonic regression.

3 The Forced Sale Discount

3.1 Static hedonic regression

Hedonic regression is a standard approach for estimating the relationship between the prices of houses and their characteristics. Our main estimating equation for measuring the forced sales discount is specified using equations such as the following for the log price, y_{ist}, of house i in census tract s in year t:

$$y_{ist} = \alpha_{st} + \beta' X_i + \lambda' F_i + \epsilon_{ist}. \tag{1}$$

Here, F_i represents measures of whether the transaction is classified as forced. For instance, in one model, it is simply an indicator if the transaction is forced, while in another model it is a vector of indicators corresponding to different types of forced sales. The terms α_{st} are census tract-year effects, which allow for house price variation over time at the census tract level. All specifications also include month dummies to control for seasonality in the housing market. X_i is a vector of house characteristics with coefficient β, and ϵ_{ist} is an error term which reflects random fluctuation in house prices. The standard errors are cluster-corrected at census tract-year cells using the procedure implemented by the *Stata* cluster command.

If F_i were randomly assigned, ordinary least squares (OLS) estimates of equation (1) would measure the average causal effects of forced sales on transaction prices. Our set of controls X_i, which is fully described in the online appendix, is unusually rich; it includes interior area, lot area, numbers of rooms, bedrooms, and bathrooms, the age of the house and its square, and dummies for recent renovation, condominiums, and winsorization of characteristics. Nonetheless, there is still a concern that forced indicators may be correlated with unobserved characteristics of the house, biasing the OLS estimates. This possibility cautions us against interpreting estimates of λ as causal. However, we suspect that unexpected forcing events such as sudden deaths are close to randomly assigned. Furthermore, if a forcing event is correlated with unobserved changes in housing characteristics that lead to lower prices, then our estimate may be interpreted as the total effect of the forced sale and the associated adverse change in unobserved housing characteristics, a point we explore in further detail below.

than unforced sales, tend to be on smaller lots and tend to be older. To make a comparison between all characteristics in a parsimonious manner, in that table, we also predict the log house price using our main hedonic regression model, equation (1), and regress this predicted price on the four forced indicators in Column (8). We find that sales that are forced by old deaths and foreclosures tend to affect houses whose characteristics would normally make them slightly cheaper than average, by about 2% and 4% respectively.

Table 2 reports our estimates of λ for three different specifications for the forced sale variable. In Panel A, the forced sale variable is an indicator if the transaction is forced. In Panel B, it is a vector of four indicators for deaths of young sellers (those who died under age 70), deaths of old sellers (those who died at age 70 or above), bankruptcy-related transactions and foreclosures. In Panel C, these four forced sale variables are interacted with dummies if there are one or two sellers. The estimates of β, the coefficients on house characteristics, are of less interest but we report them in appendix Table A.5 for the specification in Panel B. These coefficients have the expected signs and plausible magnitudes. The R^2 statistics of the specifications reported in Table 2 range from 0.72 to 0.82.

The first column of Table 2 reports results for our full sample including all housing types. When we use a single dummy for all categories of forced sales, we find a large and precisely estimated coefficient of -0.197, corresponding to a price discount of $1 - \exp(-0.197) = 18\%$.

This effect is primarily driven by foreclosure-related sales. In Panel B, when we include separate dummies for death-related sales by young and old sellers, bankruptcy-related sales, and foreclosure-related sales, we find coefficients of -0.053, -0.069, -0.035, and -0.314, respectively. The coefficient for foreclosure implies a large price discount of 27%.

In Panel C, we look separately at transactions with a single seller and with two sellers. Again, the first column reports results for all housing types. We find a much larger discount for death-related sales when the house has a single seller than when it has two sellers. In the former case the discount coefficients are -0.083 and -0.097 for young and old sellers respectively, while in the latter case they are -0.038 and -0.053. We also find a considerably larger discount for bankruptcy-related sales when there is only one seller (-0.064) than when there are two (-0.017).[8]

We have investigated the persistence of the forced sale discount by including information on the price at which each house was previously sold. We first identify the date of the most recent previous sale of each house in our transactions dataset, the price of that previous sale, and whether the previous sale was forced. We create dummy variables for previous sales that took place within the year before the current sale, one to three years before the current sale, three to five years before the current sale, and five years or more before the current sale. Then we interact the previous sales price, and dummies indicating whether the previous sale was forced, with these dummies for the timing of the previous sale. The estimates are presented in appendix Table A.10, which shows that previous sales prices do have a persistent effect, which is almost invariant to the length of time since the last sale.[9] Controlling

[8] We have explored how the estimate of the forced sales discount varies along other dimensions of our dataset. The appendix reports estimates of models where the forced sale discount varies by year (Table A.6), by the timing of the forcing event relative to the sale (Table A.7), by two subperiods 1987-1996 and 1997-2009 (Table A.8), and by geographical location in Western and Eastern Massachusetts (Table A.9).

[9] The coefficient on the previous sales price of about 0.15 implies that a 10% lower price at the time of the last sale, unexplained by the other variables in the hedonic regression, is associated with a 1.4% lower price at the time of the

for the general persistence of house prices, we do not find that forced sales have large dynamic effects. Perhaps the most interesting result is that if the previous sale was death-related, there is a modest positive effect on the subsequent sales price that roughly offsets the persistent negative effect of the death-related component of the previous sales price.

3.2 Interpreting the forced sale discount

A key challenge is to understand whether lower prices for forced sales reflect illiquidity in the housing market, or unobserved variation in fundamental characteristics of houses. For example, deaths are more common among older sellers, whose houses may be poorly maintained or unfashionably decorated. The fact that the death-related discount is increasing in the age of the seller suggests the relevance of this point. Sellers in financial difficulty may also fail to maintain their houses properly, and houses that have been foreclosed may have been vandalized while standing empty, or even in some cases vandalized by their former owners.

To shed some light on this issue, we explore how the forced sale discount varies with the timing of a sale in relation to death or bankruptcy, across housing types, and across houses whose value is concentrated in the structure or the land.

Figure 2 shows that discounts for death-related sales are relatively insensitive to the timing of the death, from 3 years before to 3 years after the sale. The somewhat larger estimate for transactions before death possibly reflects urgent sales driven by medical needs; however, when we include dummies for death-related sales more than three years before or after the date of the death (which would not be classified as forced sales), we find that these also enter the regression significantly. This confirms the suspicion that much of the estimated price effect is not directly related to the urgency of the sale, but results from unobserved poor maintenance.

The timing pattern for bankruptcy-related sales is more suggestive of a true forced-sale effect. The largest coefficient is for a sale that occurs within one year after a bankruptcy filing, and this coefficient, at -0.056, is more than twice as large as those estimated for the relatively infrequent sales that occur before bankruptcy.

In the case of foreclosures (not shown in the figure) the timing pattern is U-shaped. The coefficient is -0.308 for foreclosure-related sales within one year of foreclosure, -0.428 for sales 1 to 2 years after foreclosure, and -0.430 for sales 2 to 3 years after foreclosure. In the case of sales more than 3 years after foreclosure, the coefficient is -0.207. Since more than 85% of foreclosure-related sales occur

current sale. This persistent price effect, which is exploited by repeat-sales house price indexes (Case and Shiller 1987, 1989), could reflect unmeasured quality differentials across houses or the use of previous prices as reference prices in bargaining by sellers and buyers.

within a year of foreclosure, the deeper price discounts for the relatively small number of sales that occur with a delay of a year or more may reflect difficult market conditions that reduce the ability of a lender to dispose of a foreclosed property in a timely manner.

The right hand columns of Table 2 show how forced-sale discounts vary with housing type. Overall and foreclosure-related discounts are larger for condominiums and multi-family houses, and smaller for single-family houses. However, death-related discounts are largest for single-family houses, smaller for multi-family houses, and very small for condominiums. Since a large part of the maintenance of condominiums is handled collectively through the condominium association, and tenants in multi-family housing enforce minimum maintenance standards, this pattern is also consistent with the view that death-related discounts are related to poor home maintenance by older sellers.

To the extent that a forced sale discount reflects poor maintenance of a house, then it should be larger when the structure accounts for a greater share of the value of a property, and smaller when the land and its associated building rights account for a greater share of value. In the extreme case where a small house is sold in an expensive neighborhood as a "tear-down", there should be no maintenance-related discount at all. Thus we can measure the importance of the maintenance effect by looking at variation in the forced sale discount across houses with different hedonic characteristics.

In order to do this in a parsimonious manner, we follow a two-stage procedure. First, we estimate equation (1), the static hedonic regression of Table 2, omitting forced-sale indicators. We decompose the predicted log price of each house into components explained by the characteristics of the building, the size of the lot, and the census tract-year interaction. Next, we regress the log price of each house on the levels of these components, forced-sale indicators, and interactions between each of the forced-sale indicators and each of the value components standardized to have zero mean and unit standard deviation. The estimates are reported in Table 3.

The coefficients on forced-sale indicators in Table 3 are very similar to those reported earlier in Table 2. However there are some interesting interaction effects which imply larger or smaller discounts for forced sales of houses with atypical characteristics. For death-related sales the price discounts for all housing types, and for single-family houses, are larger when the building has greater value, consistent with the idea that older sellers maintain their houses poorly. For bankruptcy-related sales, the price discount is almost invariant to the value of the building, but is larger for houses in expensive census tract-years. For foreclosures, the price discount is larger when the building is less valuable, and is also larger for houses in low-priced census tract-years.

These results support the following broad interpretation of forced-sales discounts. Death-related discounts appear to result primarily from poor maintenance of single-family houses by older sellers, since the discounts are increasing in seller age, relatively insensitive to the timing of sales in relation

to death, large for single-family houses and very small for collectively maintained condominiums, and greater for houses with more valuable structures. There may also be an additional liquidity effect due to urgent medical expenses prior to death.

Bankruptcy-related discounts are consistent with a true liquidity effect. Bankrupt sellers aim to reduce their housing costs after bankruptcy, and the urgency of doing this is greater for houses in expensive census tracts because these houses have higher implicit rental costs. Bankruptcy-related discounts are higher for such houses, and higher when a house is sold the year after bankruptcy, but relatively insensitive to housing type.

Foreclosure-related discounts appear to be related both to the urgency of sale, and to vandalism. Foreclosed houses may have been vandalized during the transfer of ownership to mortgage lenders; and lenders sell urgently both because empty houses deliver no housing services, and because it is expensive to protect such houses against vandalism. Foreclosure-related discounts are larger in low-priced census tracts, and larger for cheaper houses. This pattern may reflect a greater threat of vandalism in bad neighborhoods, and fixed costs of protection that justify larger proportional discounts on cheaper houses.

4 Forced Sales and Neighborhood House Prices

4.1 Zipcode-level price dynamics

In this section we ask how the incidence and prices of forced sales relate to the prices of unforced sales. We begin by aggregating house prices to the zipcode-year level and examining the dynamics of zipcode-level house prices. In each zipcode in each year, we weight each transaction equally and calculate the average price of forced sales, the average price of unforced sales, and the share of forced sales. Appendix Table A.11 reports summary statistics for this dataset. Unsurprisingly, we again find that forced sales take place at lower prices. The distribution of the forced-sales share is extremely right-skewed, with a median of only 4% but a 99th percentile of 47%. We winsorize the fraction of forced sales at this level.

Table 4 presents regressions that describe the dynamics of house prices at the zipcode level. Each model has time and zipcode fixed effects.

In a preliminary regression, not reported in the table, we make no distinction between between forced and unforced sales prices. We regress price growth on lagged price growth and obtain a negative coefficient of about -0.43 with a standard error of 0.009, indicating that zipcode-level price variation is mean-reverting. This result contrasts with the price momentum, or positive serial correlation of price changes, observed in citywide, statewide, or national house price indexes (Case and Shiller, 1989).

The addition of lagged price growth leads to a modest improvement in the explanatory power of the regression relative to a model with only time effects of about 11%.

Next we separate log forced and unforced sales prices, and estimate an error-correction model for the two of them. More specifically, we estimate a first-order vector autoregression (VAR) for the change in log forced sales prices and the level of the forced sales discount, that is, the difference between log unforced and forced sales prices. This procedure is appropriate if the forced sales discount is stationary, so that log forced and unforced sales prices are cointegrated (Campbell and Shiller 1987, Engle and Granger 1987). The estimated VAR implies time-series behavior for the omitted variable, in this case the log unforced sales price.[10]

We find a strong tendency for reversal in forced sales price growth in Panel A of Table 4. Lagged forced price changes predict forced price changes with a coefficient of -0.07. In addition, a large discount of forced sales prices from unforced prices predicts that forced sales prices will increase. These two effects together explain an additional 38% of the variation in forced sales price growth relative to a model with only time dummies. The forced sales discount is mean-reverting, with a coefficient of 0.07 on its own lag. The discount also has a coefficient of 0.04 on lagged forced sales price growth, implying that the discount is more likely to narrow if it reached its previous level through a recent decline in forced sales prices; this is another manifestation of reversal in forced sales price growth. The equations for these two variables imply only very modest predictability for unforced sales prices, with negative coefficients of -0.03 on lagged forced sales prices and -0.09 on the lagged discount, and almost no improvement in the explanatory power relative to the model with only time effects.

These VAR results imply that both forced and unforced sales prices move in such a way as to narrow unusually large forced sales discounts. However, the additional explanatory power of the regression is much greater for forced sales prices than for unforced sales prices. Zipcode averages of unforced sales prices appear to be much closer to a random walk than are zipcode averages of forced sales prices. This result supports the view that on average within each zipcode, unforced sales take place at approximately efficient prices, while forced sale prices are mean-reverting because they reflect time-varying illiquidity in zipcode-level housing markets.

The variation over time in the incidence of forced sales allows us to ask whether zipcode-level house price dynamics are affected by this incidence. In Panel B of Table 4, we add the share of forced sales as a variable in the VAR system. We find that the forced sales share is highly persistent, with a coefficient of 0.60 on its own lag, and that it depresses forced sales price growth (with a coefficient

[10] If enough lags are included in the system, the implied dynamics are the same whether one omits the unforced or the forced sales price. We obtain broadly consistent results if we estimate a VAR for the change in log unforced sales prices and the level of the forced sales discount, including either one or two lags.

of -0.63) and widens the forced sales discount (with a coefficient of 0.58). Once again, this VAR implies very little predictability in the growth rate of unforced sales prices.

Finally, in Panel C, we consider the possibility that a high share of forced sales affects the dynamics of forced sales prices not only by directly predicting price changes, but by altering the coefficients on the other variables of the VAR system. We regress the forced sales share, the change in the log forced sales price, and the forced sales discount on their own lags and the interaction of the lagged forced sales share with the other two explanatory variables. We find that a high forced sales share reduces the tendency for forced sales price growth to reverse, and reduces the response of forced sales price growth to the forced sales discount. Consistent with this, a high forced sales share increases the persistence of the forced sales discount. The autoregressive coefficient for the forced sales discount increases from 0.05, in an environment with an average 6% share of forced sales, to 0.28, in an environment with a share of forced sales at the 47% winsorization point. In other words, a location with a high share of forced sales is likely to have persistently depressed forced sales prices and high forced sales discounts.

In all these specifications, we continue to find that unforced sales price growth is hard to predict. For unforced sales price growth, even the rich model estimated in Panel C adds only 5.4% of explanatory power to a model with only time dummies. The incremental explanatory power increases modestly if we add additional VAR lags, but never exceeds 15% in any of the models we have estimated. The limited predictability of zipcode-level house price movements, when sales are unforced, is a robust result in our dataset.

4.2 Local effects of foreclosures

Even though forced sales do not seem to drive large predictable movements in average unforced sales prices within the same zipcode, it is possible that there are more local effects of forced sales on neighboring houses that do not show up in data aggregated to the zipcode level. A particular concern is that houses vacated during the foreclosure process drive down neighborhood house prices. In this section we use data on the precise location of each house transaction in our dataset to try to identify such effects. Our main approach is to add variables to our hedonic regression that measure the number of foreclosures, defined as cases in which ownership of neighboring houses has been transferred to mortgage lenders, causing them to enter an urgent sales process. We find considerable evidence that foreclosures within 0.25 mile, and particularly within 0.1 mile, lower the price at which a house can be sold.

A challenge in interpreting this result is that local economic shocks, such as plant closings, may drive both house prices and foreclosures. Furthermore, foreclosures are endogenous to house prices because homeowners are more likely to default if they have negative equity, which is more likely as

house prices fall. Ideally, we would like an instrument that influences foreclosures but that does not influence house prices except through foreclosures; however, we have not been able to find such an instrument.

Instead, we compare the effects of foreclosures before and after each transaction, and the effects of extremely close foreclosures (under 0.1 mile from the target house) with those that occur further away within the 0.25 mile radius. To the extent that common economic shocks affect house prices and foreclosures within broad local areas, they should not create stronger effects of extremely local foreclosures. To the extent that house prices drive foreclosures, low prices should precede foreclosures rather than vice versa. For a foreclosure in neighborhood s in year t, our strategy compares average log house prices for all houses that transacted after the foreclosure within a 0.25 mile radius to average log house prices for all houses that transacted before the foreclosure. If there is a common shock in the neighborhood which generates an overall downward trend within this micro-geography, it will be captured by the difference between these two groups. Our main assumption is that within this small geography, a foreclosure should have differential effects on the prices of houses that are within even closer proximity. This is captured by the comparison of average log house prices for houses that transacted before and after the foreclosure within 0.10 miles. The difference between past and future foreclosure coefficients within 0.10 miles, controlling for past and future foreclosures within the far radius, gives us a difference-in-difference estimate of the causal effect of foreclosures on nearby house prices.[11]

To implement this approach, we enrich our earlier regression model by including measures of nearby foreclosures as explanatory variables. Let $N_{k,l}$ denote the number of foreclosures within geographic region $k \in \{close, far\}$ and time period $l \in \{before, after\}$. The models we report in the main text define the geographic radius for far and close to be 0.25 and 0.10 miles, respectively. Before refers to all transactions in the year prior to the sale, while after refers to all transactions in the year following the sale. The appendix reports estimates from a series of models where we vary these definitions.

The model we estimate is a variation of the following:

$$y_{ist} = \alpha_{st} + \beta' X_i + \lambda' F_i + \delta_{C,B} \cdot g_1(N_{C,B}) + \delta_{C,A} \cdot g_2(N_{C,A}) + \delta_{F,B} \cdot h_1(N_{F,B}) + \delta_{F,A} \cdot h_2(N_{F,A}) + \epsilon_{ist}, \quad (2)$$

where $g(\cdot)$ and $h(\cdot)$ are functions that allow us to parameterize the effects of multiple foreclosures. For *far*, we let $h(\cdot)$ be the sum of the number of foreclosures within 0.25 miles. As with equation (1),

[11] Our strategy was inspired by Linden and Rockoff (2008)'s study of the effect of sex offenders on house prices. Like us, they face the challenge that sex offenders are not randomly assigned to neighborhoods. They infer a causal effect to sex offender arrival by comparing house prices before and after a sex offender moves into a neighborhood, and house prices closer to the sex offender's address with those further away. The difference between house price growth in the sex offender's immediate neighborhood and house price growth in the sex offender's broader neighborhood is an estimate of the effect of the sex offender's arrival on house prices.

the standard errors are clustered at census tract-year.

Because the distribution of foreclosures is extremely right-skewed, one concern is that a few outliers dominate our estimates. We are, however, particularly interested in the effects of foreclosure waves on house prices. Our preferred specification is a piece-wise linear function where the pieces are allowed to have different slopes between the 99th and 99.5th percentile (11-17 foreclosures), between the 99.5th and 99.9th percentile (17-31 foreclosures), and above the 99.9th percentile up to the sample maximum (31-74 foreclosures). To capture this in the regression equation, we interact the $\delta_{F,B}$ and $\delta_{F,A}$ terms with indicators for these segments.[12] Because 81% of our transactions have no foreclosures within 0.25 mile during the year before sale, these tail dummies capture a meaningful fraction of the cases with foreclosures. For example, 0.01/0.19 or 5.2% of transactions with foreclosures nearby are above the 99th percentile of the foreclosure distribution.

For *close*, in the main text, we report estimates where $g(\cdot)$ is a distance-weighted sum of foreclosures where the weight is 0.1 less the distance to the foreclosure in miles, divided by 0.1. This indicator gives a weight of 1 to a foreclosure at the same location (which can occur in a condo complex), a weight of 0.5 to a foreclosure 0.05 miles or 88 yards away, and a weight of zero to a foreclosure 0.1 miles or 176 yards away. We report estimates with this weighting function as it is plausible that spillover effects of foreclosures on crime and the social cohesion of neighborhoods are extremely local, more so than common economic shocks that might drive both foreclosures and house prices.[13] The estimated impact of a foreclosure on home values at the same location as the foreclosure is given by the difference $\delta_{C,B} - \delta_{C,A}$. As with far, the close distribution is right-skewed, so we estimate a piece-wise linear model by including dummies for extreme cases (1.70-2.66, 2.66-7.33, and 7.33-64). 92% of our transactions have no foreclosures within 0.1 mile during the year before sale, so as before, the tail dummies include a meaningful fraction of the cases with extremely close foreclosures.

Table 5 reports the estimates from this model. All previous controls are included (including the indicators for forced sales) but we report the values of δ. The first two columns only utilize information

[12] To write our estimating equation, first define link function:

$$\begin{aligned} m(\boldsymbol{\delta}, f, N) = {} & \mathcal{I}\{N < N^{99.0}\} \cdot [\delta \cdot f(N)] + \mathcal{I}\{N^{99.5} \geq N \geq N^{99.0}\} \cdot [\delta \cdot f(N^{99.0}) + \delta^{99.0} \cdot (f(N) - f(N^{99.0}))] \\ & + \mathcal{I}\{N^{99.9} \geq N \geq N^{99.5}\}[\delta \cdot f(N^{99.0}) + \delta^{99.0} \cdot (f(N^{99.5}) - f(N^{99.0})) + \delta^{99.5} \cdot (f(N) - f(N^{99.5}))] \\ & + \mathcal{I}\{N \geq N^{99.9}\}[\delta \cdot f(N^{99.0}) + \delta^{99.0} \cdot (f(N^{99.5}) - f(N^{99.0})) + \delta^{99.5} \cdot (f(N^{99.9}) - f(N^{99.5})) + \delta^{99.9} \cdot (f(N) - f(N^{99.9}))] \end{aligned}$$

where $\mathcal{I}\{\cdot\}$ is an indicator function, $\boldsymbol{\delta}$ is a four-vector, $(\delta, \delta^{99.0}, \delta^{99.5}, \delta^{99.9})$, and, e.g, $N^{99.0}$ refers to the value of N at the 99.0th percentile. The estimating equation is:

$$y_{ist} = \alpha_{st} + \beta' X_i + \lambda' F_i + m(\boldsymbol{\delta_{C,B}}, g_1, N_{C,B}) + m(\boldsymbol{\delta_{C,A}}, g_2, N_{C,A}) + m(\boldsymbol{\delta_{F,B}}, h_1, N_{F,B}) + m(\boldsymbol{\delta_{F,A}}, h_2, N_{F,A}) + \epsilon_{ist}.$$

[13] Appendix Table A.17 reports specifications with alternate weighting functions including no weighting and shows that estimates reported in Table 5 are largely insensitive to choice of weighting function for multiple foreclosures.

on the number of nearby foreclosures before the sale of the house; they report $\delta_{C,B}$ and $\delta_{F,B}$ in equation (2), together with the slope coefficients for the extreme values.

In the second column, we also control for average prices of unforced sales within the 0.25 mile radius during the previous year to allow for micro-level effects within this small neighborhood. We calculate a weighted average of log prices (a geometric average price), using a linear weighting scheme that gives a weight of 0.25 less the distance to the house in miles, divided by 0.25. By contrast with the local foreclosure indicator, this is a weighted average, not a weighted sum, so it divides by the sum of the weights. We set the variable to zero in cases where no unforced transaction has occurred within 0.25 miles during the previous year, and include a dummy for these cases.

In the third and fourth columns, we add information on the number of foreclosures after the sale of the house and the average neighborhood house prices during the year after each transaction. If unobserved local shocks drive both prices and foreclosures, or if foreclosures react to prices with a lag, we would expect that future foreclosures would have at least as much explanatory power for house prices as lagged foreclosures. In columns (3) and (4) we report the difference in the coefficients and the implied standard errors: the estimate of $\delta_{F,B} - \delta_{F,A}$ is reported in the first row of the table, while the estimate of $\delta_{C,B} - \delta_{C,A}$ is reported in the second row.

The first two columns of Table 5 imply that recent neighborhood foreclosures are highly relevant for predicting the price at which a house will sell. Each foreclosure within a 0.25 mile radius of a given house lowers the predicted log price by 1.7% in column (1), or 1.1% in column (2) when we control for the average level of recent unforced sales prices in the neighborhood. Foreclosures within a 0.1 mile radius are even stronger predictors, lowering the log price of a house by 8.7% if the foreclosure is at zero distance, or 7.2% when we control for recent unforced sales prices, numbers close to those claimed recently by the Obama Administration (US Treasury 2009). In the tail of the distribution the magnitudes of these slope coefficients decrease, implying that the overall effect of nearby foreclosures is concave in the number of foreclosures. Nonetheless, this overall effect is extremely large in the tails. A house in the top 0.1% of the distribution for both variables has a price forecast that is lower by over 30% in column (1), or about 27% in column (2).[14]

The third and fourth columns of Table 5 show that recent foreclosures are stronger negative predictors of house prices than are future foreclosures. The differences between before and after coefficients, $\delta_{F,B} - \delta_{F,A}$ and $\delta_{C,B} - \delta_{C,A}$, are consistently negative. The difference $\delta_{C,B} - \delta_{C,A}$ in column (3) tells us that a foreclosure at zero distance lowers the price of a house by 2.0% more if it took place within the past year than if it will take place within the next year, controlling for the number of foreclosures within a 0.25 mile radius of the house. This can be interpreted as a difference-

[14] This calculation follows, e.g., from (-0.087)*1.70 + (-0.055)(2.66 - 1.70) + (-0.037)(7.338 - 2.661) = -0.37 log points or 31% in column (1).

in-difference estimate of the causal effect of a foreclosure on the prices of nearby houses. In column (4), we control for nearby unforced sales prices and still obtain a difference $\delta_{C,B} - \delta_{C,A}$ of 1.7%. A typical foreclosure within the 0.1 mile radius takes place at a distance of 0.05 miles; such a foreclosure gets a weight of 0.5 in the nearby foreclosure index, implying a negative spillover effect of 1.0% in column (3), and 0.85% in column (4).

What do these estimates imply about the effects of the current foreclosure wave? As a rough calculation, we have studied the effects of the actual foreclosures that took place during 2008 on all neighboring houses, whether or not these houses were actually sold. If we use the forecasting model in column (2) of Table 5, the typical foreclosure during this period lowered the price of the foreclosed house by $44,000 and the prices of neighboring houses by a total of $477,000, for a total loss in housing value of $520,000. If we use the difference-in-difference estimate from column (4) of Table 5, the typical foreclosure in 2008 lowered the price of the foreclosed house by $44,000 and the prices of neighboring houses by a total of $148,000, for a total loss of $192,000. Even this considerably smaller estimate implies that foreclosures have important negative effects on the prices of nearby houses.

Several other results about spillovers are reported in the appendix. In Table A.12, we distinguish between foreclosed properties that are already sold by the time a neighboring house is sold, and those that are still on the market. There is little difference between the estimated spillovers in these two cases. We consider whether our estimated spillovers differ between the first and current housing downturn in Massachusetts in Table A.13. When we compare estimates using data only from 1987-1996 vs. using only data from 1997-2009, our estimated spillover from the early period is marginally larger than the later period, though the difference is not statistically significant. In Table A.14, we compare Eastern and Western Massachusetts, and find precise estimates for both parts of the state, with a larger Eastern Massachusetts effect. The estimates at the 99.0th percentile are of comparable magnitudes, however.

The other dimensions we explore are alternate geographic definitions for far and close, alternate timing definitions for before and after, and alternative schemes for weighting multiple foreclosures. For all of these dimensions, the broad patterns are unchanged relative to Table 5. Interested readers can find the estimates in Tables A.15, Table A.16 and Table A.17.

In Table A.18, we investigate the role of lagged foreclosures on neighborhood house prices by enriching the specification in column (1) and (2) of Table 5 to add one further lagged year of our near and far foreclosure measures. We find that there there is a significant effect of foreclosures that happened between one and two years before a house is sold, an effect that is present even when we include controls for average house prices within the 0.25 mile neighborhood. While these estimates do not control for future foreclosures, their persistence suggests that foreclosures do not merely cause

transitory liquidity discounts on the prices of neighboring houses, but may have negative physical effects on neighborhoods which last for some time. If this is the case, it adds credibility to the concern that foreclosures reduce the ability of neighbors to refinance their mortgages, and may even drive down neighbors' home equity to the point at which they also have incentives to default. In Table A.19, we estimate spillovers separately by housing type in the top panel and interactions with value components in the bottom panel. There are two additional facts in this table. The largest estimated spillover is for condominiums. In the overall sample, there is also evidence that properties that are located in worse neighborhoods within a census tract-year experience a larger negative spillover.

Finally, we use the same strategy to estimate the spillover effect of deaths and bankruptcies. The estimates are reported in Table A.20. In contrast to foreclosures, we are unable to detect a negative spillover effect from either deaths or bankruptcies. Relatedly, we examine whether neighborhood foreclosures affect the discount at which forced transactions take place in Table A.21. The effects on bankruptcies and deaths are imprecise. However, we find that foreclosures within 0.25 mile of a house tend to increase the discount at which a foreclosed house is sold relative to comparable unforced sales, consistent with our zipcode-level finding in Table 4, but foreclosures within 0.1 mile tend to reduce that discount.[15]

Our results cannot be definitive on the causality from foreclosures to house prices, but the combination of timing effects (stronger from lagged foreclosures than from future foreclosures) and geographical effects (stronger at extremely short distances) suggests that there is reason to be concerned about spillovers from foreclosures to neighboring houses despite the reassuring zipcode-level results reported in the previous subsection.

5 Conclusion

This paper uses data on more than 1.8 million house transactions in Massachusetts to show that houses sold after foreclosure, or close in time to the death or bankruptcy of at least one seller, are sold at lower prices than other houses. The discount is particularly large for foreclosures, 27% of a house's value on average. It is smaller for death-related sales at 5-7% of value, and smaller again for bankruptcy-related sales at 3% of value.

The pricing pattern for death-related sales suggests that the discount may be due to poor maintenance, because it does not depend sensitively on the timing of the sale relative to the timing of a seller's death, is larger for deaths of older sellers, and is larger for houses where the structure accounts for a greater fraction of the value of the property. The pricing pattern for foreclosures is quite different.

[15] We also examined the sensitivity of the results in Table 5 to the inclusion of foreclosed transactions. When we remove these transactions, the implied spillover is slightly larger than when they are included.

Foreclosure discounts are larger for low-priced properties in low-priced census tracts, which suggests that foreclosing mortgage lenders face fixed costs of homeownership, probably related to vandalism, that induce them to accept absolute discounts that are proportionally larger for low-priced houses.

After aggregating to the zipcode-year level and controlling for movements in the overall level of Massachusetts house prices, we find that the prices of unforced transactions are close to a random walk, while forced sales take place at a substantial and time-varying discount. This discount is larger and more persistent when the share of forced sales is higher. These patterns suggest that most unforced transactions in residential real estate take place at efficient prices, at least relative to the general level of house prices in Massachusetts. Forced sales take place at lower prices, which one might think of as revealing a "bid price" for houses as in the finance literature on the bid-ask spread in dealer markets (e.g. Roll 1984). When many homeowners are selling urgently, the implied bid-ask spread widens for housing.

We also look for evidence that forced sales have spillover effects on the prices of local unforced sales. This question is of particular interest given the increase in the foreclosure rate in the current housing downturn (Gerardi, Shapiro, and Willen 2007, Calomiris, Longhofer, and Miles 2008). We find that foreclosures predict lower prices for houses located less than 0.25 mile, and particularly less than 0.1 mile away. Although foreclosures and prices are both endogenous variables, the fact that foreclosures lead prices at such short distances does reinforce the concern that foreclosures have negative external effects in the housing market. Our preferred estimate of the spillover effect suggests that each foreclosure that takes place 0.05 miles away lowers the price of a house by about 1%.

References

Apgar, William C., Mark Duda, and Rochelle N. Gorey, 2005, "The Municipal Cost of Foreclosures: A Chicago Case Study", unpublished report, Homeownership Preservation Foundation, Minneapolis, MN.

Benmelech, Efraim, Mark J. Garmaise, and Tobias J. Moskowitz, 2005, "Do Liquidation Values Affect Financial Contracts? Evidence from Commercial Loan Contracts and Zoning Regulation," *Quarterly Journal of Economics* 120, 1121–1154.

Calomiris, Charles W., Stanley D. Longhofer, and William Miles, 2008, "The Foreclosure-House Price Nexus: Lessons from the 2007–2008 Housing Turmoil", NBER Working Paper No. 14294.

Campbell, John Y., Stefano Giglio, and Parag Pathak, 2009, "Appendix", available online at http://kuznets.fas.harvard.edu/~campbell/papers.html.

Campbell, John Y. and Robert J. Shiller, 1987, "Cointegration and Tests of Present Value Models", *Journal of Political Economy* 95, 1062–1088.

Case, Bradford and John M. Quigley, 1991, "The Dynamics of Real Estate Prices", *Review of Economics and Statistics* 73, 50–58.

Case, Karl E. and Robert J. Shiller, 1987, "Prices of Single-Family Homes Since 1970: New Indexes for Four Cities", *New England Economic Review*, September/October, 45–56.

Case, Karl E. and Robert J. Shiller, 1989, "The Efficiency of the Market for Single-Family Homes", *American Economic Review* 79, 125–137.

Engle, Robert F. and Granger, Clive W. J., 1987, "Cointegration and Error-Correction: Representation, Estimation, and Testing", *Econometrica* 55, 251–276.

Genesove, David and Christopher J. Mayer, 1997, "Equity and Time to Sale in the Real Estate Market", *American Economic Review* 87, 255–269.

Genesove, David and Christopher J. Mayer, 2001, "Loss Aversion and Seller Behavior: Evidence from the Housing Market", *Quarterly Journal of Economics* 116, 1233–1260.

Gerardi, Kristopher, Adam Shapiro, and Paul Willen, 2007, "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures", unpublished paper, Federal Reserve Bank of Boston.

Harding, John P., Eric Rosenblatt, and Vincent W. Yao, 2008, "The Contagion Effect of Foreclosed Properties", unpublished paper, University of Connecticut.

Immergluck, Dan and Geoff Smith, 2005, "There Goes the Neighborhood: The Effect of Single-Family Mortgage Foreclosures on Property Values", unpublished paper, Woodstock Institute, Chicago, IL.

Immergluck, Dan and Geoff Smith, 2006, "The Impact of Single-Family Mortgage Foreclosures on Neighborhood Crime", *Housing Studies* 21, 851–866.

Levitt, Steven D. and Chad Syverson, 2008, "Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate", *Review of Economics and Statistics* 90, 599–611.

Linden, Leigh and Jonah E. Rockoff, 2008, "Estimates of the Impact of Crime Risk on Property Values from Megan's Laws", *American Economic Review* 98, 1103–1127.

Mayer, Christopher J., 1995, "A Model of Negotiated Sales Applied to Real Estate Auctions", *Journal of Urban Economics* 38, 1–22.

Mayer, Christopher J., 1998, "Assessing the Performance of Real Estate Auctions", *Real Estate Economics* 26, 41–66.

Meese, Richard A. and Nancy E. Wallace, 1997, "The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches", *Journal of Real Estate Finance and Economics* 14, 51–73.

Morgan, Donald P., Benjamin Iverson, and Matthew Botsch, 2008, "Seismic Effects of the Bankruptcy Reform", Federal Reserve Bank of New York Staff Report No. 358.

Roll, Richard, 1984, "A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market", *Journal of Finance* 39, 1127–1139.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens III, 2008, "Housing Externalities", NBER Working Paper No. 14369.

Shleifer, Andrei and Robert W. Vishny, 1992, "Liquidation Values and Debt Capacity: A Market Equilibrium Approach", *Journal of Finance* 47, 143–166.

Springer, Thomas M., 1996, "Single-Family Housing Transactions: Seller Motivations, Price, and Marketing Time", *Journal of Real Estate Finance and Economics* 13, 237–254.

US Treasury, 2009, *Making Home Affordable: Updated Detailed Program Description,* available online at http://treas.gov/press/releases/reports/housing_fact_sheet.pdf.

White, Michelle J., 2008, "Bankruptcy: Past Puzzles, Recent Reforms, and the Mortgage Crisis", NBER Working Paper No. 14549, forthcoming *American Law and Economics Review.*

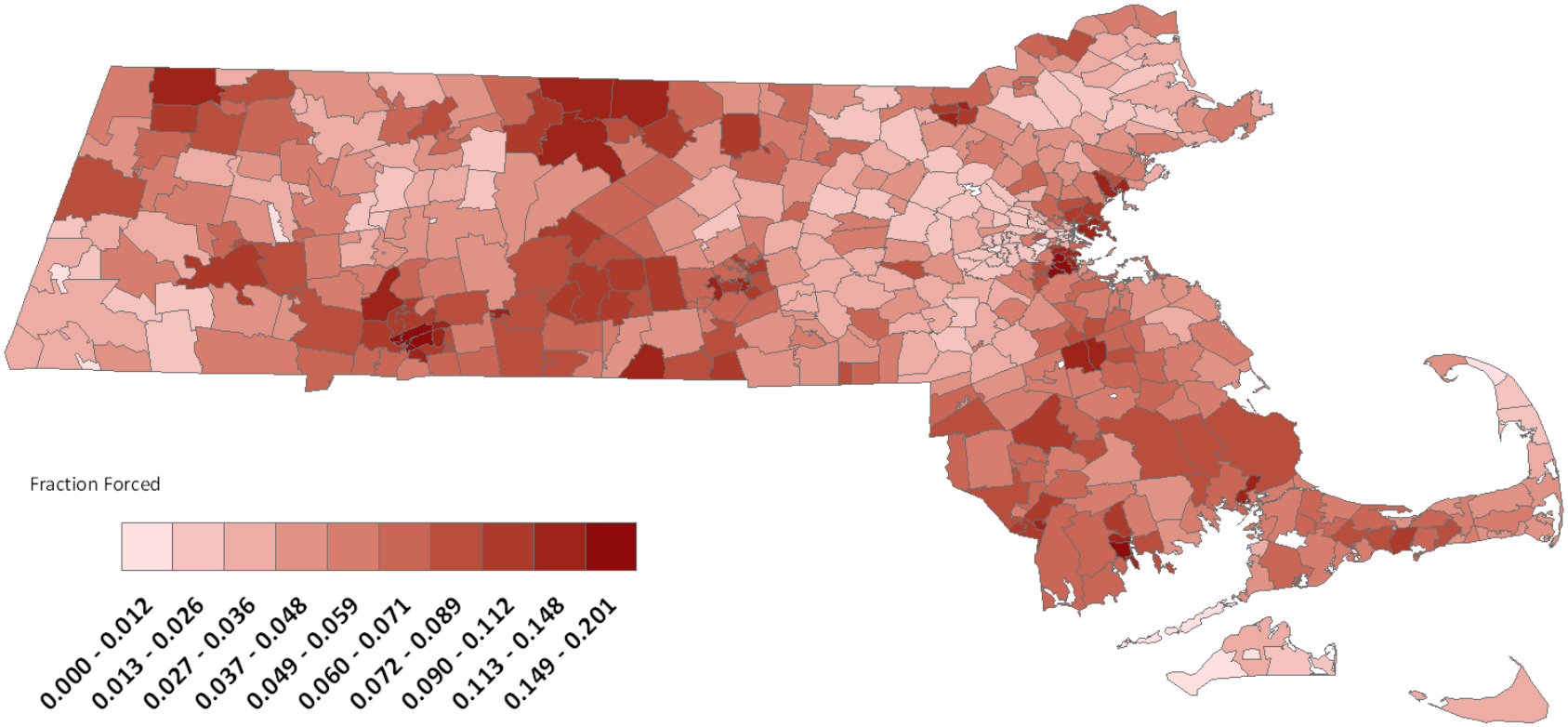

Figure 1: Geographic Distribution of Percentage of Housing Transactions that are Forced Sales by Zip Code

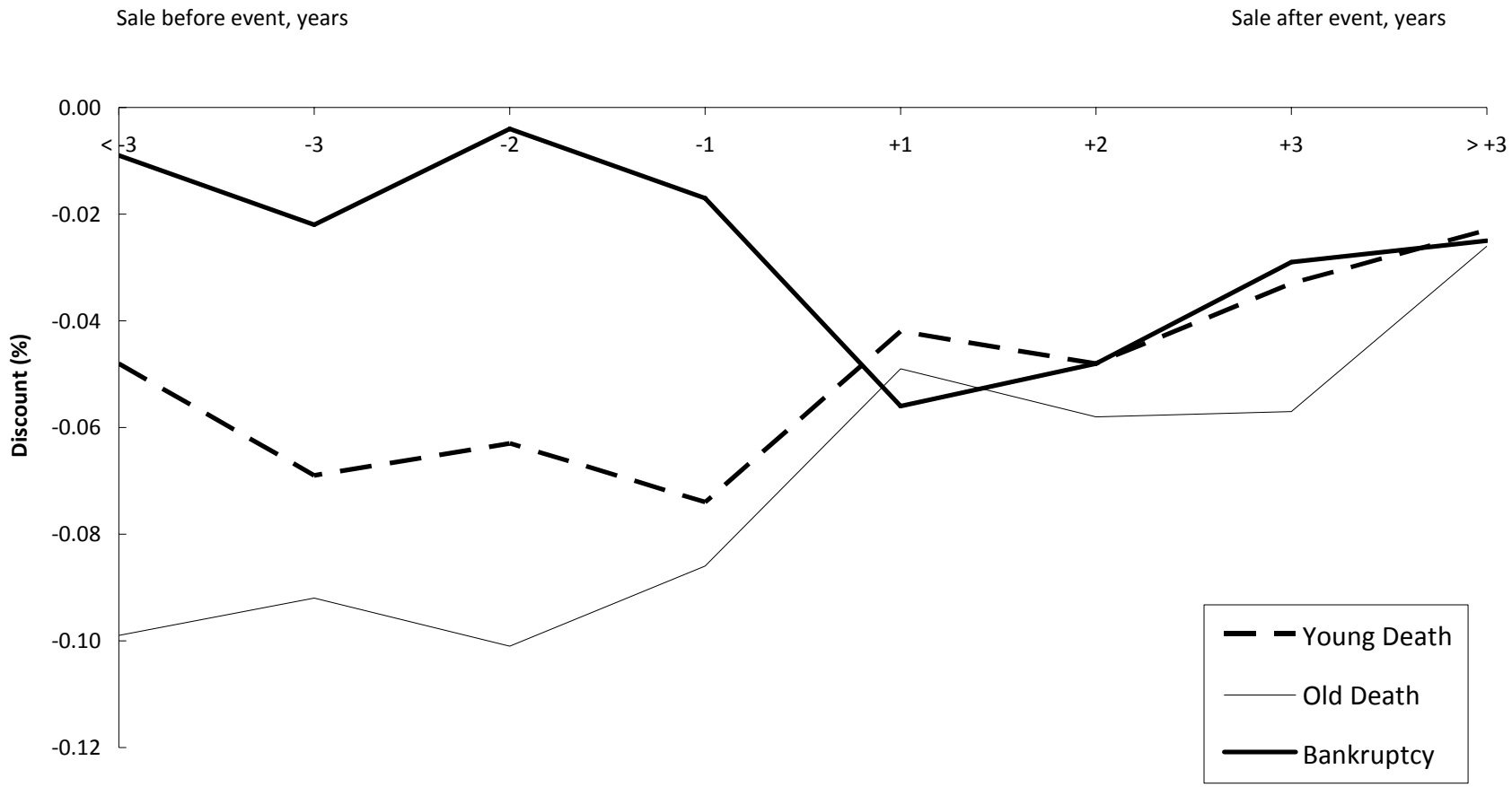

Figure 2: Forced Sales Discount and Time Between Sale and Event

Table 1 - Frequency and Timing of Forced Sales

Panel A: Number of Forced Transactions by Year

Year	Total Observations	Deaths	Bankruptcies	Foreclosures	Total Forced
1987	87,257	1.1%	-	0.0%	1.1%
1988	78,461	0.9%	-	0.0%	0.9%
1989	65,728	0.9%	-	0.3%	1.2%
1990	54,062	1.0%	-	1.1%	2.1%
1991	57,013	1.1%	0.1%	5.2%	6.4%
1992	68,471	1.2%	0.2%	8.2%	9.6%
1993	74,556	1.6%	0.3%	9.4%	11.4%
1994	81,058	1.8%	0.5%	8.3%	10.5%
1995	75,909	1.8%	0.6%	7.0%	9.3%
1996	84,046	1.6%	0.7%	4.9%	7.3%
1997	90,163	1.8%	0.8%	4.3%	6.9%
1998	99,770	1.9%	0.9%	3.0%	5.7%
1999	103,247	1.8%	1.1%	2.3%	5.2%
2000	95,036	1.9%	1.1%	1.8%	4.8%
2001	89,555	2.0%	1.2%	1.4%	4.5%
2002	92,582	2.2%	1.2%	1.2%	4.6%
2003	94,692	2.3%	1.4%	0.7%	4.5%
2004	105,630	2.5%	1.4%	0.7%	4.6%
2005	101,929	2.4%	1.3%	0.8%	4.5%
2006	86,243	2.3%	1.3%	1.6%	5.2%
2007	77,526	2.2%	0.9%	5.3%	8.4%
2008	60,483	1.9%	0.7%	14.0%	16.6%
2009(Q1)	7,976	2.1%	0.7%	25.7%	28.5%
Total	1,831,393	1.8%	0.8%	3.5%	6.1%

Panel B: Timing of Forced Transactions Relative to Forcing Event

Group	Death	Bankruptcy	Foreclosure
sale 3 yrs before event	12.9%	10.3%	
sale 2 yrs before event	15.2%	10.1%	
sale 1 yr before event	20.6%	9.5%	
sale 1 yr after event	29.1%	30.8%	85.9%
sale 2 yrs after event	14.8%	22.1%	9.1%
sale 3 yrs after event	7.4%	17.2%	1.6%

Notes: data on deaths from the Social Security Death Master file and data on bankruptcies obtained from the MA Bankruptcy Court, which begins in 1993. Panel A reports the fraction of observations classified as deaths, bankruptcies, or foreclosures each year. An observation is assigned to one of the mutually exclusive categories according to the rules described in the Online Appendix. For deaths and bankruptcies, a sale is considered forced if the sale happens 3 years either before or after the sale. For foreclosures, a sale is considered forced whenever if the sale occurs after the foreclosure auction or if the auction itself is successful. For each type of forced sale, Panel B reports how the fraction of forced sales is distributed relative to forcing event. The table represents all transactions in Massachusetts from 1987 through March 2009, with sample restrictions described in the Online Appendix. In Panel B, the remaining 3.4% of transactions in the foreclosure column represents transactions which happen more than 3 years after the foreclosure.

Table 2 - Price Discount for Forced Sales

	Full sample		*Single Family*		*Multi Family*		*Condominium*	
	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: All Forced Transactions								
Forced (-3 years;+3 years)	-0.197	(0.002)	-0.159	(0.002)	-0.236	(0.004)	-0.224	(0.005)
Panel B: Forced Transactions by Type								
Death, young seller (-3;+3)	-0.053	(0.005)	-0.068	(0.006)	-0.027	(0.014)	-0.012	(0.012)
Death, old seller (-3;+3)	-0.069	(0.002)	-0.082	(0.003)	-0.062	(0.007)	-0.017	(0.006)
Bankruptcy (-3;+3)	-0.035	(0.003)	-0.042	(0.003)	-0.020	(0.008)	-0.033	(0.007)
Foreclosure	-0.314	(0.003)	-0.260	(0.003)	-0.344	(0.005)	-0.308	(0.006)
Panel C: Forced Transactions by Number of Sellers								
One seller								
Death, young seller (-3;+3)	-0.083	(0.010)	-0.093	(0.012)	-0.057	(0.026)	0.007	(0.018)
Death, old seller (-3;+3)	-0.097	(0.004)	-0.107	(0.005)	-0.099	(0.013)	-0.025	(0.011)
Bankruptcy (-3;+3)	-0.064	(0.005)	-0.073	(0.006)	-0.024	(0.012)	-0.051	(0.010)
Two sellers								
Death, young seller (-3;+3)	-0.038	(0.006)	-0.056	(0.007)	-0.009	(0.016)	-0.022	(0.015)
Death, old seller (-3;+3)	-0.053	(0.003)	-0.070	(0.003)	-0.041	(0.008)	-0.013	(0.008)
Bankruptcy (-3;+3)	-0.017	(0.004)	-0.025	(0.004)	-0.016	(0.011)	-0.014	(0.010)
Number of Observations	1,831,393		1,187,645		202,123		441,625	

Notes: table reports estimates and standard errors, in parenthesis, of a regression of log house price on house characteristics and disaggregated forced sale indicators, on the full sample, and separately for each house type. Coefficients on house characteristics for the full sample specification are reported in the Online Appendix. Death, bankruptcy and foreclosure indicators are mutually exclusive. Young seller is defined as a seller younger than 70 at the time of death. There are 5,715 cases of young deaths and 27,134 cases of old deaths. 45% of the sample has two sellers. The regression includes census tract-year fixed effects. Standard errors are clustered at the census tract-year level.

Table 3 - Price Discounts and Value Components

	Full sample		*Single Family*		*Multi Family*		*Condominium*	
	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Main forced sale effect								
Death, young seller (-3;+3)	-0.055	(0.005)	-0.070	(0.007)	-0.023	(0.015)	-0.007	(0.012)
Death, old seller (-3;+3)	-0.071	(0.002)	-0.090	(0.003)	-0.062	(0.007)	-0.016	(0.007)
Bankruptcy (-3;+3)	-0.034	(0.003)	-0.040	(0.004)	-0.019	(0.010)	-0.033	(0.007)
Foreclosure	-0.283	(0.003)	-0.250	(0.003)	-0.281	(0.005)	-0.299	(0.007)
Forced sale effects interacted with building component								
Death, young seller (-3;+3)	-0.017	(0.007)	-0.006	(0.008)	0.032	(0.017)	0.023	(0.012)
Death, old seller (-3;+3)	-0.031	(0.004)	-0.014	(0.004)	0.004	(0.008)	-0.013	(0.008)
Bankruptcy (-3;+3)	-0.007	(0.004)	-0.003	(0.004)	0.007	(0.009)	-0.001	(0.008)
Foreclosure	0.041	(0.003)	0.039	(0.003)	-0.007	(0.005)	0.035	(0.006)
Forced sale effects interacted with lotsize component								
Death, young seller (-3;+3)	0.000	(0.006)	-0.005	(0.006)	0.015	(0.016)	-0.027	(0.015)
Death, old seller (-3;+3)	0.000	(0.003)	-0.019	(0.003)	0.016	(0.007)	0.009	(0.011)
Bankruptcy (-3;+3)	-0.006	(0.004)	-0.011	(0.004)	-0.014	(0.010)	0.003	(0.009)
Foreclosure	-0.008	(0.003)	0.004	(0.003)	0.017	(0.005)	-0.014	(0.007)
Forced sale effects interacted with tract-year component								
Death, young seller (-3;+3)	-0.011	(0.006)	-0.005	(0.006)	-0.013	(0.018)	-0.026	(0.013)
Death, old seller (-3;+3)	-0.004	(0.002)	0.001	(0.003)	-0.006	(0.008)	-0.018	(0.007)
Bankruptcy (-3;+3)	-0.023	(0.004)	-0.021	(0.004)	-0.015	(0.011)	-0.006	(0.009)
Foreclosure	0.036	(0.003)	-0.005	(0.003)	0.082	(0.005)	-0.002	(0.005)

Notes: table reports estimates and standard errors, in parenthesis, from regressions of log price with indicators of forced sales, plus the interactions with standardized components of the value of the house. These components are first obtained from a regression of log price on the house characteristics of the regression in Panel B of Table 2. The predicted price is decomposed into the components explained by the value of the building, the size of the lot, and the census tract-year interaction. The reported estimates are from a second regression of log price on the forced sales dummies interacted with the components described above, standardized to zero mean and unit variance. This regression includes the other regressors of Table 2. Standard errros are clustered at the census tract-year level.

Table 4 - VAR for Neighborhood House Prices

	sf_t		Δpf_t		pu_t-pf_t		Δpu_t	
	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err	Estimate	Std Err
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: VAR				
Δpf_{t-1}	-	-	-0.072	(0.012)	0.038	(0.012)	-0.033	(0.004)
pu_{t-1}-pf_{t-1}	-	-	0.840	(0.016)	0.067	(0.016)	-0.093	(0.006)
Δ Adj R^2			0.384		0.144		0.016	
				Panel B: VAR with lagged forced share				
sf_{t-1}	0.596	(0.012)	-0.633	(0.070)	0.575	(0.070)	-0.057	(0.027)
Δpf_{t-1}	-0.004	(0.002)	-0.067	(0.012)	0.034	(0.012)	-0.033	(0.004)
pu_{t-1}-pf_{t-1}	0.006	(0.003)	0.857	(0.016)	0.051	(0.016)	-0.092	(0.006)
Δ Adj R^2	0.368		0.390		0.152		0.016	
				Panel C: VAR with lagged forced share interactions				
sf_{t-1}	0.548	(0.016)	0.166	(0.097)	0.320	(0.098)	0.486	(0.037)
Δpf_{t-1}	0.011	(0.003)	-0.044	(0.015)	0.041	(0.015)	-0.003	(0.006)
pu_{t-1}-pf_{t-1}	0.007	(0.003)	0.984	(0.020)	0.018	(0.020)	0.002	(0.008)
sf_{t-1} x Δpf_{t-1}	-0.263	(0.026)	-0.082	(0.153)	-0.202	(0.154)	-0.284	(0.057)
sf_{t-1} x (pu_{t-1}-pf_{t-1})	0.038	(0.031)	-1.989	(0.181)	0.565	(0.181)	-1.424	(0.068)
Δ Adj R^2	0.373		0.400		0.153		0.054	

Notes: table reports estimates and standard errors, in parenthesis, from vector auto-regressions (VAR) of percentage change in average forced and unforced house prices at the zipcode-year level. *pf* is the average price of forced sales, *pu* the average price of unforced sales, and *sf* the share of forced sales in each zipcode at time t. Each specification includes neighborhood and time fixed effects. The number of observations in each regression is 7,254. The reported "Δ Adj R^2" is the difference between the adjusted R^2 of the full model and the adjusted R^2 of the model with only time dummies.

Table 5 - Estimates of Spillover Effects of Foreclosures

	Using only Foreclosures Before Transaction: Before [$\delta_{F,B}$ and $\delta_{C,B}$]		Estimated Difference in Coefficients: Before - After [($\delta_{F,B}$ - $\delta_{F,A}$) and ($\delta_{C,B}$ - $\delta_{C,A}$)]	
	(1)	(2)	(3)	(4)
Slope: far (δ_F)	-0.017	-0.011	-0.006	-0.003
	(0.001)	(0.001)	(0.005)	(0.001)
Slope: close (δ_C)	-0.087	-0.072	-0.020	-0.017
	(0.003)	(0.003)	(0.001)	(0.003)
Slope at 99.0: far ($\delta_F^{99.0}$)	0.002	-0.000	-0.011	-0.007
	(0.002)	(0.002)	(0.004)	(0.003)
Slope at 99.0: close ($\delta_C^{99.0}$)	-0.055	-0.050	-0.048	-0.043
	(0.012)	(0.011)	(0.017)	(0.014)
Slope at 99.5: far ($\delta_F^{99.5}$)	-0.004	-0.003	-0.008	-0.005
	(0.002)	(0.002)	(0.009)	(0.002)
Slope at 99.5: close ($\delta_C^{99.5}$)	-0.037	-0.030	-0.031	-0.027
	(0.007)	(0.006)	(0.003)	(0.008)
Slope at 99.9: far ($\delta_F^{99.9}$)	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.003)	(0.002)
Slope at 99.9: close ($\delta_C^{99.9}$)	-0.009	-0.005	0.001	0.002
	(0.003)	(0.003)	(0.002)	(0.004)
Additional controls:				
Average price, before		0.248		0.180
		(0.002)		(0.002)
Average price, after				0.184
				(0.002)
No transaction before indicator		2.992		2.168
		(0.028)		(0.022)
No transaction after indicator				2.244
				(0.022)

Notes: table reports estimates and standard errors, in parenthesis, from regressions of log price on the unweighted number of foreclosures in the 0.25mi area around the house sold (variable *far*), and the linearly weighted number of foreclosures in the 0.1mi area (variable *close*), for the year before and after the sale. The effect is specified as piecewise linear in the intervals (0-99th pct), (99th-99.5th), (99.5th-99.9th), (99.9th-max), with the estimated coefficients reported. Columns (1) and (2) are models which only use foreclosures that happened before each sale. The reported estimates are the slope coefficients of each part of the piece-wise linear function. Columns (3) and (4) include the foreclosures before and after the sale. The reported estimates are the difference in the estimates for each piece of the piece-wise linear function, between the effect of foreclosures before and after the transaction. Columns (2) and (4) also include the distance-weighted average log price of neighboring houses (0.25mi), in the year before and after the sale, and an indicator for the cases where there are no transactions in the neighborhood, in that time frame. Each model includes the house and forced sale characteristics of Table 2, Panel B. The cutoff points in the piece-wise linear function for close are: 1.696 (99th percentile), 2.661 (99.5 percentile), and 7.338 (99.9th percentile). For the far variable, the cutoff points are: 11 (99th percentile), 17 (99.5th percentile), and 31 (99.9th percentile). Standard errors are clustered at the census tract-year level.