How Do Foreclosures Exacerbate Housing Downturns?

Adam M. Guren and Timothy J. McQuade*
Boston University and Stanford University
December 3, 2014

Abstract

We present a dynamic search model in which foreclosures exacerbate housing busts and delay the housing market’s recovery. By raising the seller to buyer ratio and making buyers more selective, foreclosures freeze the market for non-foreclosures and reduce price and sales volume. Because negative equity is necessary for default, foreclosures can cause price-default spirals that amplify an initial shock. To quantitatively assess these channels, the model is calibrated to the recent bust. The estimated amplification is significant: foreclosures exacerbated aggregate price declines by 60 percent and non-foreclosure price declines by 24 percent. Furthermore, policies that slow foreclosures can be counterproductive.

Keywords: Housing Prices & Dynamics, Foreclosures, Search, Great Recession.

*guren@bu.edu, tmcquade@stanford.edu. We would like to thank John Campbell, Edward Glaeser, and Jeremy Stein for outstanding advice and Pat Bayer, Karl Case, Raj Chetty, Darrell Duffie, Emmanuel Farhi, Erik Hurst, Lawrence Katz, Greg Mankiw, Chris Mayer, Monika Piazzesi, Andrei Shleifer, Alp Simsek, Johannes Stroebel and seminar participants at Harvard University, Stanford University, the NBER Summer Institute, the Penn Search and Matching Workshop, and the Greater Boston Urban and Real Estate Economics Seminar for helpful comments. We would like to acknowledge CoreLogic and particularly Kathryn Dobbyn for providing data and answering numerous questions.
Foreclosures have been one of the dominant features of the recent housing downturn. From 2006 through 2013, approximately eight percent of the owner-occupied housing stock experienced a foreclosure. Although the wave of foreclosures has subsided, foreclosures remain at elevated levels, and understanding the role of foreclosures in housing downturns remains an important part of reformulating housing policy.

The behavior of the housing market concurrent with the wave of foreclosures is shown in Figure 1. Real Estate Owned (REO) sales—sales of foreclosed homes by mortgage servicers—made up over 20 percent of existing home sales nationally for four years. Sales of retail (non-foreclosure) volume fell 65 percent as time to sale rose. Prices dropped considerably, with aggregate price indices plunging a third and retail prices falling by a quarter.

This paper presents a model in which foreclosures have important general equilibrium effects that can explain much of the behavior of housing markets during the bust, particularly in the hardest-hit areas. By raising the number of sellers and reducing the number of buyers, by making buyers more choosy, and by changing the composition of houses that sell, foreclosure sales freeze up the market for retail sales and reduce both price and sales. Furthermore, the effects of foreclosures can be amplified considerably because price declines induce more default which creates further price declines, generating a feedback loop. A quantitative calibration suggests that these effects can be large: foreclosures exacerbated aggregate price declines nationwide in the recent bust by approximately 60 percent and retail price declines by 24 percent.

We first present a dynamic stochastic equilibrium search model of the housing market with random moving shocks, undirected search, idiosyncratic house valuations, Nash bargaining over price, and endogenous conversion of owner-occupied homes to rental homes. The model builds on a literature on search frictions in the housing market, notably Wheaton (1990), Williams (1995), Krainer (2001), Novy-Marx (2009), Ngai and Tenreyo (2014), and Head et al. (2014). We use this structural framework to calibrate the key parameters in our model to pre-downturn moments.

We introduce foreclosures by assuming they differ from normal sales in two key ways: REO sellers have higher holding costs and individuals who are foreclosed upon cannot immediately buy a new house. Foreclosures dry up the market for normal sales and reduce volume and price through three main effects on market equilibrium. First, because foreclosed upon homeowners are locked out of the market, foreclosures reduce the likelihood that a seller will meet a buyer in the market through a “market tightness
Figure 1: The Role of Foreclosures in the Housing Downturn

Notes: All data is seasonally adjusted national-level data from CoreLogic as described in the online appendix. The grey bars in panels B and C show the periods in which the new homebuyer tax credit applied. In panel C, all sales counts are unsmoothed and normalized by the maximum monthly existing home sales while each price index is normalized by its separate maximum value.
This channel emphasizes that foreclosures do not expand supply but rather reduce demand. The effect is somewhat blunted by endogenous conversion of owner-occupied homes into rental homes. Second, the presence of distressed sellers increases the outside option to transacting for buyers, who have an elevated probability of being matched with a distressed seller next period and consequently become more selective. This novel “choosey buyer effect” endogenizes the degree of substitutability between bank and retail sales. Third, there is a mechanical “compositional effect” as the average sale looks more like a distressed sale.

We embed this model of the housing market into a richer framework featuring mortgage default, in which underwater homeowners default if hit by a liquidity shock and are otherwise “locked in” to their current house when other reasons to move arise. In conjunction with the effect of foreclosures on prices, default amplifies the effects of negative shocks to prices: an initial shock that reduces prices puts some homeowners under water and triggers foreclosures, which cause more price declines and in turn further default. Lock-in of underwater homeowners also impacts market equilibrium by keeping potential buyers and sellers out of the market, increasing the share of listings which are distressed.

We fit the model with default to match the cross-section of downturns across metropolitan areas from 2006 to 2013. Motivated by the fact that the size of the bust is highly correlated with the size of the preceding boom, we simulate a downturn triggered by a permanent decline in home valuations proportional to the size of the preceding run-up in prices. Because sales fell considerably even in metropolitan areas with a small price boom and bust, we also include a nation-wide temporary slowdown in moving shocks. We calibrate the fraction of the boom that is permanently lost in the bust and the size of the national slowdown in moving shocks to minimize the distance between the model simulation of the downturn and data on the downturn for 97 of the 100 largest metropolitan areas. The calibrated model fits a number of moments that are not used in calibration despite the fact that most parameters are calibrated to pre-downturn moments.

The model is used to quantitatively evaluate the extent to which foreclosures have exacerbated the ongoing housing bust and to assess housing policy. Qualitatively, the model explains two key features of the data: that areas that had a larger boom had a more-than-proportionately larger bust, and that this effect was concentrated in areas with a combination of a larger boom and a large number of homeowners with low
equity at the beginning of the crisis. Quantitatively, we can use the model to simulate a counterfactual crisis without foreclosures to show that the effects of foreclosures exacerbated the aggregate price decline in the downturn by approximately 60 percent in the average CBSA (or account for 38 percent of the decline) and exacerbated the price declines for retail sellers by over 24 percent (or account for 20 percent of the decline).

Finally, the model can be used to analyze the equilibrium effects of a number of foreclosure policies. We focus on a policy that the model is particularly well-suited to assess: slowing down the rate at which foreclosures occur. We find that absent substantial debt forgiveness by banks in light of growing processing times or a substantial thaw in credit markets when defaults are spread over time, such a policy can be counter-productive because the benefits of having fewer foreclosures for sale at any given time are dominated by a lengthening of the foreclosure crisis.

The remainder of the paper is structured as follows. Before presenting the model, Section 1 presents facts about the downturn across metropolitan areas. Section 2 introduces our basic model of the housing market and Section 3 estimates the parameters using pre-downturn moments. Section 4 builds on the housing market model by adding foreclosures and default. Section 5 calibrates the full model to the recent downturn, assesses the model fit, and discusses its implications and counterfactuals. Section 6 considers foreclosure policy, and Section 7 concludes.

1 Empirical Facts

The national aggregate time series for price, volume, foreclosures, and REO share presented in Figure 1 mask substantial heterogeneity across metropolitan areas. To illustrate this, Figure 2 shows price and volume for four of the hardest-hit Core Based Statistical Areas (CBSAs). In Las Vegas, for instance, prices fell nearly 60 percent, and the REO share was as high as 75 percent.

To provide a more systematic analysis of the heterogeneity of the bust across cities and to motivate, calibrate, and test the model, we use a proprietary data set provided by CoreLogic supplemented by data from the United States Census. CoreLogic provides monthly data for 2000 to 2013 for the nation as a whole and 99 of the 100 largest CBSAs. The data set includes a repeat sales house price index, a house price index for retail sales only, sales counts for REOs and retail sales, and estimates
Figure 2: Price and Sales in Selected MSAs With High Levels of Foreclosure

Notes: All data is seasonally adjusted CBSA-level data from CoreLogic as described in the online appendix. Sales are smoothed using a moving average and normalized by the maximum monthly existing home sales, while each price index is normalized by its maximum value.

of quantiles of the LTV distribution. We seasonally adjust the CoreLogic data and smooth the sales count series using a moving average. A complete description of the data and summary statistics are in the online appendix.

The best predictor of the size of the bust is the size of the preceding boom. Figure 3 plots the change in log price from 2003 to 2006 against the change in log price from each market’s peak to its trough. There is a strong downward relationship, which motivates a key feature of our model: the shock that causes the bust is a fall in home valuations that is assumed to be proportional to the size of the preceding boom.

Figure 4 provides a similar analysis for sales, comparing the size of the boom in prices from 2003 to 2006 against the peak-to-trough decline in sales of non-REO homes. Although the size of the run-up is related to the size of the decline in sales,
sales fell significantly even in CBSAs that had virtually no boom. To explain this fact, our model incorporates a nation-wide slowdown in the hazard rate of moving in addition to downward shock to home valuations that is proportional to the size of the preceding boom.

Figure 3 also reveals a more subtle fact: places that had a larger boom had a more-than-proportional larger bust. While a linear relationship between log boom size and log bust size has an r-squared of .62, adding a quadratic term that allows for larger busts in places with larger booms increases the r-squared to .68. The curvature can

Note: Scatter plot of seasonally adjusted data from CoreLogic along with quadratic regression line that excludes CBSAs in greater Detroit which busted without a boom. The data is described in the online appendix. Each data point is an CBSA and is color coded to indicate in which quartile the CBSA falls when CBSAs are sorted by the share of homes with over 80 percent LTV in 2006. Although the highest LTV MSAs had almost no boom and no bust (e.g. Indianapolis), the CBSAs below the best fit line tend to be MSAs with a large share of homeowners with high LTVs in 2006 (third quartile).
Figure 4: Price Boom vs. Sales Decline Across MSAs

Note: Scatter plot of seasonally adjusted and smoothed data from CoreLogic. The data is described in the online appendix. Although the decline in sales volume was closely related to the size of the price run-up, even those CBSAs with almost no price run-up experienced a large decline in sales.

be seen in the best-fit line in Figure 3.¹ We argue that by exacerbating the downturn in the hardest-hit areas, foreclosures can explain why places with larger booms had disproportionately larger busts.

This explanation has an important corollary: because default is predominantly caused by negative equity, a larger bust should occur not only in places with a larger initial downward shock to prices, but in particular in locations with a combination of a large shock and a large fraction of houses with high LTVs—and thus close to default—prior to the bust. To provide suggestive evidence that this prediction is borne out in the data, the points in Figure 3 are color-coded by quartiles of share of homeowners in the CBSA with over 80 percent LTV in 2006. While the largest high LTV shares occurred in places that did not have a bust—home values were not

¹For the best fit line, the regressions that follow, and the model calibration, we exclude two outlier CBSAs in greater Detroit which had a large bust without a preceding boom so that the non-linearity is not overstated. All results are robust to including these two CBSAs.
inflated in 2006, so the denominator of LTV was lowest in these locations—one can see that the majority of CBSAs substantially below the quadratic trend line were in the upper end of the high LTV share distribution (red triangles in the figure).

To formally investigate whether the interaction of many households with high LTVs and a big preceding boom is correlated with a deep downturn, we estimate regressions of the form:

\[
Y = \beta_0 + \beta_1 \max \Delta_{03-06} \log (P) + \beta_2 [\Delta_{03-06} \log (P)]^2 + \beta_3 (Z \max \text{Share LTV} > 80\%) + \beta_4 (\Delta_{03-06} \log (P) \times Z \text{ LTV} > 80\%) + \varepsilon.
\]

Z represents a z-score (\(z(x) = \frac{x - \text{mean}}{\text{stddev}}\)) and the outcome variable \(Y\) is either the maximum change in log price, the maximum change in log retail prices, the maximum change in log retail sales, the maximum REO share, or the fraction of houses that experience a foreclosure. This regression is similar in spirit to Lamont and Stein (1999), who show that prices are more sensitive to income shocks in cities with a larger share of high LTV households.

The regression results are shown in Table 1, with summary statistics and robustness checks in the online appendix. The fourth row shows the key result: the interaction term between the size of the run-up and the share of high-LTV homeowners is significantly negative for price, retail price, and sales volume and significantly positive for the REO share of volume and the fraction of the housing stock that is foreclosed upon. This is consistent with foreclosures playing an important role in the downturn.

2 Housing Market Model

We begin by constructing a model of the housing market during normal times, when there is a negligible amount of default. This model is used to estimate the structural parameters of the model using pre-downturn data and is the foundation for our model with foreclosures.

We consider a Diamond-Mortensen-Pissarides-style general equilibrium search model of the housing market. Search frictions play an important role in housing markets: houses are illiquid, most households own one house and move infrequently, buyers and sellers are largely atomistic, and search is costly, time consuming, and random.
Table 1: Cross-MSA Regressions on the Impact of the Size of the Bubble and Its Interaction With High LTV Share

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th>$\Delta \log (P)_{03-06}$</th>
<th>$\Delta \log (P_{\text{Retail}})$</th>
<th>$\Delta \log (\text{Sales}_{\text{Retail}})$</th>
<th>$\frac{\text{Sales}<em>{\text{REO}}}{\text{Sales}</em>{\text{Existing}}}$</th>
<th>% Foreclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \log (\text{Price})_{03-06}$</td>
<td>0.251</td>
<td>0.092</td>
<td>-1.618</td>
<td>-0.632</td>
<td>-0.362</td>
</tr>
<tr>
<td></td>
<td>(0.414)</td>
<td>(0.321)</td>
<td>(0.626)**</td>
<td>(0.404)</td>
<td>(0.139)**</td>
</tr>
<tr>
<td>$\Delta \log (\text{Price})^2_{03-06}$</td>
<td>-1.886</td>
<td>-1.537</td>
<td>0.045</td>
<td>1.837</td>
<td>0.988</td>
</tr>
<tr>
<td></td>
<td>(0.498)***</td>
<td>(0.382)***</td>
<td>(0.741)</td>
<td>(0.492)***</td>
<td>(0.171)***</td>
</tr>
<tr>
<td>$Z \text{ Share LTV} &gt; 80%$</td>
<td>0.054</td>
<td>0.061</td>
<td>-0.048</td>
<td>-0.005</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>(0.029)*</td>
<td>(0.024)**</td>
<td>(0.048)</td>
<td>(0.028)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>$\Delta \log (P) \times Z \text{ LTV} &gt; 80%$</td>
<td>-0.310</td>
<td>-0.336</td>
<td>-0.231</td>
<td>0.265</td>
<td>0.235</td>
</tr>
<tr>
<td></td>
<td>(0.123)**</td>
<td>(0.113)***</td>
<td>(0.138)*</td>
<td>(0.110)**</td>
<td>(0.054)***</td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.690</td>
<td>0.716</td>
<td>0.611</td>
<td>0.443</td>
<td>0.651</td>
</tr>
<tr>
<td>$N$</td>
<td>97</td>
<td>97</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>

Notes: * = 10% Significance, ** = 5% Significance *** = 1% significance. All standard errors are robust to heteroskedasticity. Each column shows estimates of (1) with the constant suppressed. All data is from CoreLogic and described in the online appendix. Data is for 97 largest CBSAs excluding Syracuse NY, Detroit MI, Warren MI, and, for columns 3-5, Birmingham AL. See the online appendix for robustness and regressions including southeast Michigan.
Figure 5: Schematic Representation of Housing Market Model

Time is discrete and the discount factor is $\beta$. There are a unit mass of individuals and a unit mass of houses. The model is thus a closed system with a fixed population.

The setup of the model is illustrated schematically in Figure 5 and the notation employed is summarized in Table 2. In the main text we focus on the value functions and defer the laws of motion that formalize Figure 5 to the online appendix. To simplify the analysis, we assume no default, which is approximately the case when prices are stable or rising. At any given time, mass $l_0(t)$ of individuals are homeowners. Homeowners randomly experience shocks with probability $\gamma$ that represent changes in tastes and life events that induce them to leave their house as in Krainer (2001) and Ngai and Tenreyro (2014). We assume for now that these shocks occur at a constant rate and that only individuals who receive a moving shock search for houses.

An individual who receives a moving shock enters the housing market as both a buyer and a seller. Buyers must rent a home while searching at the equilibrium rent $r(t)$. As in Head (2014), each period sellers have the option of listing their house for sale, which incurs a flow cost of $m_n$, or renting their house for $r(t)$ for one period. We assume that a given unit of housing provides rental services for $\zeta$ renter households, with $\zeta < 1$ so that renters occupy less square footage than owner-occupants as in the data. The rental market is competitive and the supply consists of those owner-occupants.

---

2 Allowing default in steady state complicates the analysis but does not substantially change the results.
Table 2: Variables in Housing Market Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endogenous Variables</strong></td>
<td></td>
</tr>
<tr>
<td>$h$</td>
<td>Stochastic Match Quality $\sim F_t(h)$</td>
</tr>
<tr>
<td>$h_n$</td>
<td>Cutoff $h$ for trade to occur</td>
</tr>
<tr>
<td>$S^B_{n,h}$, $S^S_{n,h}$</td>
<td>Buyer/seller surplus of seller given match quality $h$</td>
</tr>
<tr>
<td>$p_{n,h}$</td>
<td>Price given match quality $h$</td>
</tr>
<tr>
<td>$r$</td>
<td>Per-period rent</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Market tightness (buyer/seller ratio)</td>
</tr>
<tr>
<td>$q_s(\mu)$, $q_b(\mu)$</td>
<td>Prob seller meets buyer, buyer meets seller</td>
</tr>
<tr>
<td>$l_0$</td>
<td>Mass of homeowners</td>
</tr>
<tr>
<td>$v_b$, $v_n$</td>
<td>Masses of buyers and sellers</td>
</tr>
<tr>
<td><strong>Value Functions</strong></td>
<td></td>
</tr>
<tr>
<td>$V_h$</td>
<td>Value of owning home with match quality $h$</td>
</tr>
<tr>
<td>$V_n$</td>
<td>Value of listing house for sale</td>
</tr>
<tr>
<td>$\bar{V}_n$</td>
<td>Value of vacant house</td>
</tr>
<tr>
<td>$B$</td>
<td>Value of buyer</td>
</tr>
<tr>
<td><strong>Parameters</strong></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>Discount factor</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Probability of moving shock</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Seller’s Nash bargaining weight</td>
</tr>
<tr>
<td>$\zeta$</td>
<td>Fraction of house sqft that renter occupies</td>
</tr>
<tr>
<td>$m_n$</td>
<td>Flow cost of selling</td>
</tr>
</tbody>
</table>

occupied homes being rented out plus a permanent stock of rental homes of mass $v_{rs}$.

As in Ngai and Tenreyro (2014), we assume that the buyer and seller act as independent agents. This means that there is no interaction between the buyer’s problem or bargaining game and the seller’s, and there is no structure placed on whether an individual buys or sells first. This assumption is not innocuous, as an interaction between the liquidity of homes and the decision whether to buy or sell first can increase volatility (Anenberg and Bayer, 2013), but we make it to focus our analysis on the effects of foreclosures.

Buyers and sellers in the housing market are matched randomly each period according to a standard fixed-search-intensity constant-returns-to-scale matching function. Letting $v_b(t)$ and $v_n(t)$ be the masses of buyers and listed homes in the market at time $t$ and defining market tightness $\mu(t)$ as to the ratio of buyers to listed homes...
The probability a seller meets a buyer \( q_s \) and the probability a buyer meets a seller \( q_b \) can then both be written as functions of \( \mu \).

When matched, the buyer draws a valuation for the house in the form of a flow utility \( h \), derived from a distribution \( F_t(h) \) which is stochastic and time-varying. These valuations are common knowledge between the buyer and seller, and prices are determined by generalized Nash bargaining. If the buyer and seller decide to transact, the seller leaves the market and the buyer becomes a homeowner in \( l_0 \) deriving flow utility \( h \) from the home until they receive a moving shock. If not, the buyer and seller each return to the market to be matched next period.

Let \( V_h(t) \) be the value of being in a house with match quality \( h \) at time \( t \), \( \tilde{V}_n(t) \) the value of a seller before deciding whether to rent or sell their house, \( V_n(t) \) the value of listing a house for sale, and \( B(t) \) be the value of being a buyer. \( V_h(t) \) is equal to the flow payoff plus the discounted expected continuation value:

\[
V_h(t) = h + \beta \left\{ \gamma \left[ E_t \tilde{V}_n(t+1) + E_t B(t+1) \right] + (1 - \gamma) E_t V_h(t+1) \right\}. \tag{2}
\]

Denote the total match surplus when a buyer meets a seller and draws a match quality \( h \) at time \( t \) by \( S_{n,h}(t) \), the buyer’s surplus by \( S_{B,n,h}(t) \), and the seller’s by \( S_{S,n,h}(t) \), with \( S_{n,h}(t) = S_{B,n,h}(t) + S_{S,n,h}(t) \). Let the price of the house if it is sold be \( p_{n,h}(t) \). The buyer’s surplus is equal to the value of being in the house minus the price and their outside option of staying in the market:

\[
S_{B,n,h}(t) = V_h(t) - p_{n,h}(t) + r(t) - \beta E_t B(t+1) \tag{3}
\]

The seller’s surplus is equal to the price minus their outside option of staying in the market:

\[
S_{S,n,h}(t) = p_{n,h}(t) - m_n - \beta E_t \tilde{V}_n(t+1). \tag{4}
\]

Prices are set by generalized Nash bargaining with weight \( \theta \) for the seller.

Because utility is linear and house valuations are purely idiosyncratic, a match will result in a transaction if \( h \) is above a threshold value, corresponding to zero total surplus and denoted by \( h_n(t) \):

\[
V_{h_n(t)}(t) = m_n - r(t) + \beta \left( E_t B(t+1) + E_t \tilde{V}_n(t+1) \right). \tag{5}
\]

We can then define the remaining value functions. The value of listing a house
for sale is equal to the flow payoff plus the discounted continuation value plus the expected surplus of a transaction times the probability a transaction occurs. Because sellers who list meet buyers with probability \( q_s(\mu(t)) \) and transactions occur with probability \( 1 - F_t(h_n(t)) \), \( V_n \) is defined by:

\[
V_n(t) = m_n + \beta E_t \tilde{V}_n(t + 1) + q_s(\mu(t))(1 - F_t(h_n(t))) E_t [S_{n,h}^S(t) | h \geq h_n(t)] .
\] (6)

The value of being a buyer is defined similarly:

\[
B(t) = -r(t) + \beta E_t B(t + 1) + q_b(\mu(t))(1 - F_t(h_n(t))) E_t [S_{n,h}^B(t) | h \geq h_n(t)] .
\] (7)

The conditional expectation of the surplus given that a transaction occurs appears repeatedly in the value functions. This quantity can be simplified as in Ngai and Tenreyro (2014) by using (2) together with (3) and (4) to write total surplus as:

\[
S_{n,h}(t) = V_h(t) - V_{h_n(t)}(t) = \frac{h - h_n(t)}{1 - \beta (1 - \gamma)} .
\] (8)

The conditional expectation is

\[
E_t [S_{n,h}(t) | h \geq h_n(t)] = \frac{E_t [h - h_n(t) | h \geq h_n(t)]}{1 - \beta (1 - \gamma)} .
\] (9)

Prices can be backed out by using Nash bargaining along with the definitions of the surpluses and (8) to get:

\[
p_{n,h}(t) = \frac{\theta (h - h_n(t))}{1 - \beta (1 - \gamma)} + m_n + \beta E_t \tilde{V}_n(t + 1) .
\] (10)

This pricing equation is intuitive. The first term contains \( h - h_n(t) \), which is a sufficient statistic for the surplus generated by the match as shown by Shimer and Werning (2007). As \( \theta \) increases, more of the total surplus is appropriated to the seller in the form of a higher price. This must be normalized by \( 1 - \beta (1 - \gamma) \), the effective discount rate of a homeowner. The final two terms represent the value of being a seller next period, which is the seller’s outside option. These terms form the minimum price at which a sale can occur, so that all heterogeneity in prices comes from the distribution of \( h \) above the cutoff \( h_m(t) \).

Finally, the value of a vacant home reflects the option to either list or rent the
home, and in equilibrium sellers are indifferent between renting and listing so that:

\[ \tilde{V}_n (t) = r (t) + \beta E_t \tilde{V}_n (t + 1) = V_n (t). \]

### 3 Calibration to Pre-Downturn Moments

Because our model cannot be solved analytically, we turn to numerical simulation. This requires parameterizing some of the model primitives and selecting parameters.

For the matching technology, we use a standard Cobb-Douglas matching function so that the number of matches when there are \( b \) buyers and \( s \) sellers is \( \chi b^\xi s^{1-\xi} \). The probability a seller meets a buyer when market tightness is \( \mu \) is given by \( q_s (\mu) = \frac{\chi b^\xi s^{1-\xi}}{s} = \chi \mu^\xi \), and the probability a buyer meets a seller is \( q_b (\mu) = \frac{\chi b^\xi s^{1-\xi}}{b} = \chi \mu^\xi. \)

We parameterize the distribution of idiosyncratic valuations \( F_t (\cdot) \) as an exponential distribution with parameter \( \lambda \) shifted by a parameter \( a \), which represents the aggregate valuation of homes. We choose the exponential because the memoryless property neutralizes the effects of difficult-to-measure properties of the tail thickness of the distribution of idiosyncratic tastes. This implies that the expectation of surplus conditional on a sale is fixed because \( E_t [h - h_n (t) | h \geq h_n (t)] = \frac{1}{\chi} \), so all movements in average prices work through \( V_n (t + 1) \).

To generate housing cycles, we shock \( a \), which determines the average housing valuation. This shock is chosen as a simple approximation to more realistic shocks, such as shocks to local income (Head et al., 2014) or changes in the availability of credit that induce growth in demand for owner-occupied housing (Guren, 2013), which would have very similar effects in our model. Because house prices mean revert over 5-year horizons (Glaeser et al. 2014),\(^3\) we assume that \( a \) follows a mean-reverting stochastic process with persistence \( \rho \) and independent and identically distributed normal shocks:

\[ a (t) = \rho a (t - 1) + (1 - \rho) \tilde{a} + \varepsilon (t), \quad \varepsilon (t) \sim N \left( 0, \sigma_\varepsilon^2 \right). \]

Having parameterized the model, we turn to calibrating individual parameters. Several of the parameters correspond directly to parameters in the literature or statistics that can be readily computed from aggregate data and are set to match these.

---

\(^3\)House prices also exhibit momentum, that is autocorrelation in returns, over horizons up to three years (Guren, 2013). Our model cannot account for this feature of the data.
targets. We set the annual discount rate to five percent. With each period corresponding to a week, the weekly discount factor is $\beta = 1 - .05/52$. We set the probability of moving houses in a given week to fit a median tenure for owner occupants of approximately nine years from the American Housing Survey from 1997 to 2005, so $\gamma = .08/52$. We also use the American Housing Survey to set the fraction of an owner-occupied house’s floor space occupied by a renter, $\zeta$, to be .6 to reflect the average fraction of square footage per person and lot size occupied by renters who moved in the past year relative to owner occupants.

The elasticity of the matching function is $\xi = .84$ from Genesove and Han (2012), who use National Association of Realtors surveys to estimate the contact elasticity for sellers with respect to the buyer-to-seller ratio. The constant in the matching function, $\chi$, is chosen to be .5 to make sure the probability of matching falls on $[0, 1]$. Because this is a normalization, the results are robust to alternate choices of $\chi$.

The final parameter we choose to directly match an exogenous moment is the persistence of the shock to home valuations $\rho$. We let this moment equal the persistence of local income shocks estimated by Glaeser et al. (2014), who find an annual AR(1) coefficient of 0.885, or 0.998 weekly.

The remaining parameters do not map directly to a parameter in the data or literature and are set to minimize the distance between several empirical moments and their counterparts in simulated data. These parameters are: the flow costs of being a normal and REO seller, $m_n$ and $m_d$, sellers’ Nash bargaining weight $\theta$, the shape parameter for the distribution of idiosyncratic valuations $\lambda$, the average long-run idiosyncratic valuation $\bar{a}$, the standard deviation of shocks to valuations $\sigma_\epsilon$, and the rental stock $v_{rs}$. We target seven moments to fit these seven parameters: the average price of a house, the average REO discount, time on the market for retail houses and REO houses, the ratio of buyer to seller time on the market, and the volatility of prices and volume.\footnote{Because we want to set as many parameters as possible using pre-downturn moments, we choose $m_d$, the flow cost of selling for an REO seller introduced in Section 4, through the same procedure. We are able to simulate moments for the REO discount and REO time on the market by considering a perturbation with an infinitesimal number of foreclosures.} The target values for the moments are summarized\footnote{Heuristically, these seven moments jointly determine the seven parameters because each of the moments is predominantly determined by one of the parameters. The average price is primarily determined by the average long-run valuation for houses $\bar{a}$, while the volatility of price is primarily determined by the volatility of the shock $\sigma_\epsilon$. The REO price discount is primarily determined by the ratio of REO to non-REO seller flow costs $m_d/m_n$, and given this ratio of flow costs the relative time on the market of REO and retail homes is determined by $\lambda$, which controls the variance idiosyncratic}.
in Table 3 and detailed in the online appendix.\footnote{One parameter, the REO discount, merits additional discussion. A number of studies have attempted to estimate the REO discount using pre-crisis data and have generated estimates ranging between 10 percent (Clauretie and Denshvary, 2009) and 20 percent (Campbell et al., 2011), with most estimates near 20 percent (see survey in Clauretie and Denshvary). These studies, generally using a repeat-sales or hedonic methodology, suffer from potential bias form unobserved quality differences. If foreclosed properties are improved between the REO sale and its next listing, estimates of the REO discount will be overstated. The size of the bias is difficult to assess, so we take a conservative approach by choosing 12.5 percent as our baseline target REO discount and reporting robustness checks using 10 percent and 15 percent REO discounts in the online appendix.}

To minimize the distance between the model and the data, for each set of parameters, we simulate 100 series of 20 years and calculate the percent difference between the mean of each moment in the simulated data and the target moments. We search for the minimum of the sum of squared percent differences weighting each moment equally, starting the minimization algorithm at 25 different start values. We choose the best fit as our parameter vector.

The calibrated model matches the target moments quite closely as summarized in Table 3. The parameter values are listed in Table 4, with the parameters set exogenously in the right column and the parameters set through simulation in the left column. One parameter worth noting is the seller’s share of the match surplus, valuations. Given $\lambda$, the average time on the market for normal homes is mainly driven by the flow cost of being a seller $m_n$. The ratio of buyer to seller time on the market is determined by steady state market tightness, which is principally determined by the rental stock $v_r$. Finally, the volatility of volume relative to price is determined primarily by the size of the seller surplus, which is controlled by the seller bargaining weight $\theta$ conditional on the other parameters as described in the text below.
Table 4: Parameter Values Calibrated to Pre-Downturn Moments

<table>
<thead>
<tr>
<th>Param</th>
<th>Value</th>
<th>Units</th>
<th>Param</th>
<th>Value</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_n$</td>
<td>0.005</td>
<td>Thousands of $</td>
<td>$\beta$</td>
<td>$1 - \frac{0.08}{52}$</td>
<td>Weekly Rate</td>
</tr>
<tr>
<td>$m_d$</td>
<td>$-0.146$</td>
<td>Thousands of $</td>
<td>$\gamma$</td>
<td>$\frac{0.08}{52}$</td>
<td>Weekly Rate</td>
</tr>
<tr>
<td>$\theta$</td>
<td>0.053</td>
<td></td>
<td>$\chi$</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>$\lambda$</td>
<td>3.587</td>
<td>Thousands of $</td>
<td>$\xi$</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>$\bar{a}$</td>
<td>3.150</td>
<td>Thousands of $</td>
<td>$\zeta$</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>$\sigma_\varepsilon$</td>
<td>0.092</td>
<td></td>
<td>$\rho$</td>
<td>0.998</td>
<td>Weekly Persistence</td>
</tr>
<tr>
<td>$v_{rs}$</td>
<td>0.014</td>
<td>Mass (Housing stock = 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

which is 5.3 percent. This value is set to a low level because volume is substantially more volatile than price in the data. Intuitively, price is equal to a markup over the seller’s outside option, which itself is a discounted sum of probability of sale times expected seller surplus. For price to be relatively insensitive to probability of sale and thus sales volume, the seller surplus and therefore $\theta$ must be small.

4 Model With Foreclosures

Using the model introduced in Section 2 as a foundation, we introduce foreclosures and default. The full model is then used for a quantitative analysis of the role of foreclosures in the housing bust.

4.1 Introducing Foreclosures

We assume foreclosures are different from non-foreclosures in two simple ways that capture many of the salient features of foreclosures.\footnote{Foreclosures may have other effects. They may cause negative externalities on neighboring properties due to physical damage, the presence of a vacant home, or crime. Campbell et al. (2011) show that such effects are small and highly localized. There may also be buyer heterogeneity with respect to their willingness to purchase a foreclosure, generating an additional channel through which the REO discount widens as non-natural buyers purchase foreclosures. Finally, foreclosures may cause banks to reduce credit supply, as shown theoretically by Chatterjee and Eyigungor (2011).}

First, foreclosed properties are sold by mortgage servicers, which we model has having a higher per period holding cost than a normal seller ($m_d < m_n$) for several reasons. Most importantly, servicers have substantial balance sheet concerns because they must make payments to security holders until a foreclosure liquidates, and they must also assume the costs of pursuing the foreclosure, securing, renovating, and

\begin{align*}
\text{First, foreclosed properties are sold by mortgage servicers, which we model has having a higher per period holding cost than a normal seller ($m_d < m_n$) for several reasons. Most importantly, servicers have substantial balance sheet concerns because they must make payments to security holders until a foreclosure liquidates, and they must also assume the costs of pursuing the foreclosure, securing, renovating, and}
\end{align*}
maintaining the house, and selling the property. Even though they are paid additional fees to compensate for the costs of foreclosure and are repaid when the foreclosed property sells, the servicer’s effective return is far lower than its opportunity cost of capital. In a detailed review of the mortgage servicing industry, Theologides (2010) concludes that “...once a loan is delinquent, there is no extraordinary reward that would justify exceptional efforts to return the loan to current status or achieve a lower-than-anticipated loss.” Owner-occupants also have lower costs of maintenance and security, and REO sellers usually leave a property vacant and thus forgo rental income or flow utility from the property.\(^8\)

Second, we assume that homeowners who experience a foreclosure are prevented from buying for a period of time and must rent in the interim. Foreclosure dramatically reduces a borrower’s credit score, and many banks, the GSEs, and the FHA require buyers to wait several years after a foreclosure before they are eligible for a mortgage. In practice, Molloy and Shan (2013) use credit report data to show that households that experience a foreclosure start are 55-65 percentage points less likely to have a mortgage two years after a foreclosure start. For simplicity, we assume that each period, renters who experienced a foreclosure become eligible to buy with probability \(\sigma\). As with buyers who are renting a house while they search, we assume that renters who experienced a foreclosure occupy less square footage so that each unit of housing provides rental services for \(\zeta < 1\) renters.

The setup alters the baseline housing market model only slightly. One can define the Bellman equation of an REO servicer using the same Bellman equation as for a normal seller but replacing \(m_d\) by \(m_n\):\(^9\)

\[
V_d(t) = m_d + \beta E_t V_d(t + 1) + q_s(\mu(t)) (1 - F_t(h_d(t))) E_t [S^S_{d,h}(t) | h \geq h_d(t)]
\]

\[
S^S_{d,h}(t) = S^S_{d,h}(t) = p_{d,h}(t) - m_d - \beta E_t V_d(t + 1).
\]

The buyer’s value function has to be adjusted for the fact that there are two types of sellers operating in the market. Conditional on matching with a seller, we assume that

---

\(^8\)An implicit assumption is that no deep-pocketed and patient market maker buys from distressed sellers and holds the property until a suitable buyer is found. This is likely due to agency problems and high transactions costs. We also assume that the temporary conversion of for-sale properties to rentals is only done by retail sellers.

\(^9\)Unlike REO sellers, normal sellers also act as buyers in the market at the same time. This difference is not captured in the seller value functions due to our assumption that buying and selling are quasi-independent activities.
the probability it is a particular type of seller equals that seller’s current proportion in the population, which we denote by \( r_j(t) \) for \( j = n, d \), so that the buyer’s value function can be written as:

\[
B(t) = -r(t) + \beta E_t B(t + 1) + \sum_{j=n,d} q_b(\mu(t)) r_j(t) (1 - F_t(h_j(t))) E_t [S_{j,h}^R(t) | h \geq h_j(t)],
\]

where the buyer’s surplus trading with an REO seller \( S_{d,h}^R(t) \) is defined similarly to Section 2. The price of a foreclosure can also be written as before:

\[
p_{d,h}(t) = \frac{\theta (h - h_d(t))}{1 - \beta (1 - \gamma)} + m_d + \beta E_t V_d(t + 1).
\]

Foreclosures also alter the laws of motion of the system. Define \( f_t \) as the number of foreclosures at time \( t \) and \( v_f \) as the number of individuals who have been foreclosed upon and are renting. Beyond introducing a law of motion for \( v_f \), one need only alter the previous laws of motion to adjust the inflow of normal buyers to account for an additional \( \sigma v_f \) of new buyers and adjust the inflow of new renters to account for the fact that \( \zeta v_f \) foreclosed-upon homeowners have to rent. Before endogenizing \( f_t \) by modeling default, we discuss how foreclosures affect the housing market’s equilibrium.

### 4.2 General Equilibrium Effects of Foreclosures

The qualitative results in our model are caused by the interaction of three different effects: the “market tightness effect,” the “choosey buyer effect,” and the “compositional effect.”

First, because foreclosed homeowners are locked out of the housing market as renters and only gradually flow back into being buyers, foreclosures reduce market tightness \( \mu(t) \). This decreases the probability a seller meets a buyer in a given period and triggers endogenous responses as each party’s outside option to transacting changes, altering the bargaining and sale cutoffs. In particular, the decline in market tightness reduces the value of being a seller and raises the value of being a buyer. These movements in outside options incentivize sellers to transact faster, weakening their bargaining positions and leading to lower prices. Conversely, buyers are more willing to walk away from a deal, strengthening their bargaining position which also leads to declines in prices. The market tightness effect unambiguously leads to
declines in the unconditional probability of sale, but has an ambiguous and limited effect on the probability of sale conditional on a buyer and seller meeting. The effect is stronger for REO sellers who have a higher opportunity cost of not meeting a buyer, causing the REO discount to grow. This effect is partially offset by reduction in supply cause by increased conversion of owner-occupied housing to rental space to meet the increased rental demand by foreclosed individuals. The market tightness effect demonstrates that an important element of foreclosures is a reduction in demand relative to supply, as in a typical market a move creates a buyer and a seller, while foreclosures create an immediate bank seller but a buyer only when the foreclosed upon individual’s credit improves. This contrasts with a traditional view that foreclosures are a shift out in supply rather than a reduction in demand.

Second, the value of being a buyer rises because the buyer’s outside option to transacting, which is walking away and resampling from the distribution of sellers next period, is improved by the prospect of finding an REO seller who will give a particularly good deal. This works through an increase in the REO share of listings, \( r_d \), which puts a larger weight on the REO term in (11). This term is larger because REO sellers are more likely to transact both in and out of steady state. The resulting increase in buyers’ outside options lead buyers to demand a lower price from sellers in order to be willing to transact, and in equilibrium buyers walk away from more sales in the retail market, freezing up the retail market.

The choosey buyer effect is new to the literature and formalizes folk wisdom in housing markets that foreclosures empower buyers and cause them to wait for a particularly favorable transaction.\(^\text{10}\) Albrecht et al. (2007, 2014) introduce motivated sellers into a search model, but focus on steady-state matching patterns (e.g. whether a high type buyer can match with a low type seller) and asymmetric information regarding seller type. Duffie et al. (2007) consider a liquidity shock similar to our foreclosure shock, but a transaction occurs whenever an illiquid owner meets a liquid

\( ^{10}\)For instance, The New York Times reported that “before the recession, people simply looked for a house to buy...now they are on a quest for perfection at the perfect price,” with one real estate agent adding that “this is the fallout from all the foreclosures: buyers think that anyone who is selling must be desperate. They walk in with the bravado of, ‘The world’s coming to an end, and I want a perfect place’” (“Housing Market Slows as Buyers Get Picky” June 16, 2010). The Wall Street Journal provides similar anecdotal evidence, writing that price declines “have left many sellers unable or unwilling to lower their prices. Meanwhile, buyers remain gun shy about agreeing to any purchase without getting a deep discount. That dynamic has fueled buyers’ appetites for bank-owned foreclosures” (“Buyer’s Market? Stressed Sellers Say Not So Fast” April 25, 2011).
buyer, and so while there are market tightness effects their model does not have a choosey buyer effect. We expect that choosey buyer effects may arise in other frictional asset markets with idiosyncratic valuations.

The market tightness effect and choosey buyer effect are mutually reinforcing. As discussed above, the market tightness effect is more pronounced for REO sellers. Because the value of being an REO seller falls by more, REO sellers become even more likely to sell relative to non-REO sellers during the downturn. This sweetens the prospect of being matched with an REO seller next period, amplifying the choosey buyer effect.

Finally, a greater share of REO sales makes the average sale look more like an REO, which sells faster and at a lower price both in and out of steady state. Foreclosures thus cause a mechanical compositional effect that affects sales-weighted averages such as total sales and the aggregate price index.

### 4.3 Liquidity-Driven Default

To endogenize foreclosure, we extend our model of the housing market to include default. Doing so introduces a feedback loop that amplifies the effect of an initial decline in house prices: as prices fall, more homeowners find themselves with negative equity and default, which causes more foreclosures and price declines. Negative equity is a necessary but not sufficient condition for foreclosure because homeowners with negative equity are not able to avoid default by selling their house (Foote et al., 2008).

Figure 6 schematically illustrates the model with default, and Table 5 summarizes the additional parameters. Importantly, the extended model with foreclosures only alters the mechanisms through which homeowners default and enter the housing market, so the housing market component of the model is exactly as in Section 2. Consequently, the Bellman equations and equilibrium conditions in Section 2 remain the same and only the laws of motion that formalize Figure 6 change. These laws of motion are relegated to the online appendix.

We model default as resulting from liquidity shocks that cause homeowners with negative equity to be unable to afford their mortgage payments, the so-called “double trigger” model of mortgage default. While “ruthless” or “strategic default” by borrowers has occurred, there is a consensus in the literature that strategic default
accounts for a very small fraction of mortgage defaults.\textsuperscript{11} To keep the model tractable and maintain a focus on housing market dynamics, we thus do not model strategic default, nor do we model the strategic decision of the bank to foreclose, modify the loan, rent to the foreclosed-upon homeowner, or pursue a short sale, which are options that were not widely used until late in the crisis.\textsuperscript{12}

Rather than modeling mortgage choice, refinancing, and prepayment decisions and tracking how the loan balance distribution deforms dynamically, we consider a population of mass $l_1$ of potential defaulters who have an exogenous distribution of loan balances $L$ defined by a fixed CDF $G(L)$ immediately before the bust begins.\textsuperscript{13}

\textsuperscript{11}For instance, Gerardi et al. (2013) use the PSID to estimate that only 13.9 percent of defaulters had enough liquid assets to make one monthly mortgage payment. Similarly, Bhutta et al. (2010) estimate that the median non-prime borrower does not strategically default until their equity falls to negative 67 percent, so even among non-prime borrowers in Arizona, California, Florida, and Nevada who purchased homes with 100 percent financing at the height of the bubble—80 percent of whom defaulted within 3 years—over 80 percent of the defaults were caused by income shocks. Consequently, the largest estimate of the share of defaults that are strategic is 15 to 20 percent (from Experian Oliver-Wyman). See also Elul et al. (2010) and Foote et al. (2008).

\textsuperscript{12}Modeling short sales and their effect on market equilibrium is an important topic for future research.

\textsuperscript{13}We are agnostic as to the source of the initial loan balance distribution and leave this unmodeled.
Table 5: Variables Used In Extended Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L )</td>
<td>Loan Balance ( \sim G(L) )</td>
</tr>
<tr>
<td>( w )</td>
<td>Mass of locked in homeowners</td>
</tr>
<tr>
<td>( f )</td>
<td>Mass of homeowners in foreclosure</td>
</tr>
<tr>
<td>( \gamma_f )</td>
<td>Probability of a liquidity shock</td>
</tr>
</tbody>
</table>

This takes advantage of the fact that refinancing and prepayment are rare in a downturn and that the upper tail of the loan balance distribution that is at risk of default pays down principal slowly. We also assume that population of homeowners who successfully transact during the downturn, which has mass \( l_0 \), are no longer at risk of defaulting because lenders are more conservative in their lending decisions in a bust. These simplifying assumptions imply that the population at risk of defaulting has a fixed loan balance distribution throughout the downturn, which reduces the state space considerably.

Homeowners who are underwater default when they get an income shock, which occurs with probability \( \gamma_f \), while traditional moving shocks which we call taste shocks continue to occur at rate \( \gamma \). Liquidity and taste shocks have different effects depending on the equity position of the homeowner. Homeowners with \( L \leq V_n(t) \) have positive equity and enter the housing market as a buyer and seller after either shock. Homeowners with \( L > V_n(t) \) have negative equity net of moving costs and default if they experience an income shock because they cannot pay their mortgage or sell their house. Defaulters enter the foreclosure process. Although in practice foreclosure is not immediate and some loans in the foreclosure process do “cure” before they are foreclosed upon, until we consider foreclosure policy in Section 6 we assume that foreclosure occurs immediately. We also assume that income shocks are a surprise, so an underwater homeowner expecting an income shock cannot list their house with the hope of getting a high-enough price that they can pay off their loan before the bank forecloses. While this may happen infrequently, it is unlikely that a desperate seller would receive such a high price.

Finally, homeowners with negative equity who receive a taste shock would like to move but are “locked in” their current house until prices to rise to the point that they

\( G(L) \) is fixed over time because principal is paid down slowly, particularly by those in the upper tail of the loan balance distribution who are relevant for the size of the feedback loop.
have positive equity.\textsuperscript{14} We assume that such homeowners make accommodations and thus do not immediately move when they reach positive equity, but instead wait for another taste or income shock. Lock-in further reduces sales in markets with many foreclosures and strengthens the choosey buyer effect by increasing the proportion of REO listings.

5 Cross-City Quantitative Analysis

In order to conduct counterfactuals to quantitatively assess the role of foreclosures in the crisis and to test the model’s performance in explaining the bust, we fit the model to recent downturn national and cross-city data.

5.1 Additional Calibration Targets

A few parameters are new to the model of default and are set to match the institutions and housing market on the eve of the downturn in 2006. We set the probability that a foreclosed-upon homeowner returns to owner occupancy $\sigma$ so that the average foreclosed-upon homeowner is out of the market for three years. Most lenders require one to seven years to pass after a foreclosure to be eligible for another mortgage. For instance, Veterans Administration loans require two years, Federal Housing Administration loans three years, and Fannie Mae and Freddie Mac required five years prior to 2011 and now require seven years, although reductions are allowed based on circumstances. We choose three years to fall in the middle of the range of waiting periods. We set the probability of income shocks $\gamma_I = 8.6\%$ annually using national data from CoreLogic on the incidence of foreclosure starts for houses with negative equity as described in the online appendix.

We fit the loan balance distribution to the geographic unit that we simulate, either a particular metropolitan area or the average for the nation as a whole. To do so, we parameterize the loan balance distribution $G(L)$ as a flexible beta distribution with parameters $b_a$ and $b_b$ scaled to have support on $[0, V_n]$ where $V_n$ is the value of being a normal seller in the initial stochastic steady state so that no default occurs prior to the downturn. We choose $b_a$ and $b_b$ to minimize the distance between the

\textsuperscript{14}Formally, define $w(t)$ as the mass of individuals who are locked in at time $t$. The distribution of loan balances in $w(t)$ will be the same as $G(L)$ truncated below at $V_n(t)$.\textsuperscript{25}
beta distribution and the empirical loan balance distribution in 2006 as described in the online appendix. The empirical distribution comes from proprietary estimates by CoreLogic, who report seven quantiles of the combined loan-to-value distribution for active mortgages in 2006 computed from public records and CoreLogic’s valuation models.\(^\text{15}\)

### 5.2 Fitting the Model to the Downturn

We introduce shocks to start the downturn motivated by two facts presented in Section 1: the size of the preceding price run-up is the single best predictor of the size of the ensuing downturn and sales fell nationally by a substantial amount, likely due to tightened credit.\(^\text{16}\) Consequently, we model the bust both as a combination of a shock to the local housing stock’s valuation \(a\) and a nation-wide shock to the probability of a moving shock \(\gamma\).\(^\text{17}\)

Given the empirical evidence, we model the shock to home valuations \(a\) to generate a price decline proportional to the log price gain from 2003 to 2006:

\[
\Delta \log p_{\text{permanent}} = -\eta \Delta \log p_{2003-2006}.
\]

By contrast, the shock to the probability of moving represents temporary credit conditions and so we model it as a one-time shock to \(\gamma\) of \(\Delta \gamma\) that gradually reverts to its initial value of \(\bar{\gamma}\) according to:

\[
\gamma(t) = \rho \gamma(t-1) + (1-\rho) \bar{\gamma} + \Delta \gamma,
\]

where \(\rho\) is the persistence of income shocks used in Section 3 and \(\bar{\gamma}\) is calibrated to the value of \(\gamma\) in Section 3.\(^\text{18}\)

\(^{15}\)Because our model concerns the entire owner-occupied housing stock and not just houses with an active mortgage, we supplement the CoreLogic data with the Census’ estimates of the fraction of owner-occupied houses with a mortgage from the 2005-2007 American Community Surveys.

\(^{16}\)As documented by the Federal Reserve Senior Loan Officer Survey on Bank Lending Practices, between 60 to 70 percent of banks tightened lending standards on an annual basis during the peak of the housing crisis. See http://research.stlouisfed.org/fred2/series/DRTSCLCC.

\(^{17}\)While the flow utility of housing clearly did not fall overnight, the source of the initial drop in prices is immaterial to our results, and the reduction in \(a\) should be seen as a stand-in for a number of factors that have depressed home prices, from changes in credit availability to belief disagreements to irrational exuberance.

\(^{18}\)We choose the same persistence because local income is a major contributor to the credit-worthiness of individuals who are considering moving.
With a permanent decline in $a$ and a constant hazard of an income shock, all individuals whose loan balance is above the new steady-state price level will eventually default. This does not seem realistic—many homeowners will eventually pay down their mortgage and avoid default. Rather than modeling the dynamic deformation of the $G(\cdot)$ distribution over longer horizons, we instead assume that after 5 years the hazard of income shocks $\gamma_I$ gradually subsides over the course of a year.\footnote{Formally, after 5 years $\gamma_I$ falls by 5\% of its previous value every month, taking roughly a year to settle at zero. We choose 5 years to reflect a prolonged bust.}

Our formulation of the shocks that hit the economy to trigger the downturn leaves us with two parameters that control the simulated bust across metropolitan areas: the national shock to moving hazards $\Delta \gamma$ and the fraction of the boom that is permanently lost $\eta$. We choose these parameters to minimize the distance between the model and the data for the cross-section of booms and busts experienced by 97 CBSAs.

Specifically, we choose $\eta$ and $\Delta \gamma$ to minimize the sum of squared differences between the model and the data for two equally-weighted moments: the peak-to-trough log price decline for each metropolitan area’s aggregate price index and the peak-to-trough decline in retail sales.\footnote{The price decline is predominantly determined by $\eta$ and the volume decline by $\Delta \gamma$. We choose retail sales rather than total sales because the speed with which foreclosures hit the market as REOs depended on local regulations. This can cause total volume to decline significantly more than the difference between the decline in retail sales and the increase in REO sales because the decline in retail sales occurs before REO sales ramp up. In our model, the decline in retail sales and the expansion of REO sales occurs concurrently.} To do so, we simulate the downturn for 97 of the 100 largest CBSAs and choose the parameter vector that minimizes the sum of squared residuals.

For each CBSA, we calculate the policy functions and simulate the impulse responses to the shocks to $a$ and $\gamma$, under the assumption of no further stochastic shocks to $a$. We start each CBSA in its initial stochastic steady state, introduce the two shocks in period zero, and then consider the return of the model to its new steady state. This is done because our assumptions about an at-risk population of potential defaulters at the beginning of the crisis preclude stochastic shocks to $a$ that increase prices above their original level. Furthermore, the log-linear approximation to our model would be poor with a large permanent shock to $a$, and assuming a single initial shock allows us to use more accurate solution techniques.
5.3 Model Fit

The calibration procedure yields a unique optimum, with values of $\eta = 0.632$ and $\Delta \gamma = 0.410$. These parameters imply that roughly 63 percent of the price gains between 2003 and 2006 were permanently lost and that the frequency of moving shocks for homeowners fell to approximately 41 percent of their pre-downturn level.

Figure 7 evaluates the model fit visually by plotting the simulated results against empirical data for the 97 CBSAs (black dots) and a simulation of the nation as a whole (red X). The panels show comparisons for six different metrics: the maximum log change in the aggregate price index, the maximum log change in the retail price index, the maximum log change in retail sales, the maximum change in log retail price, the maximum change in log REO sales relative to the peak level of non-REO sales, the REO share of sales, and the fraction of the housing stock that is foreclosed upon. In each figure, the 45-degree line is drawn in to represent a perfect match between the model and the data. Results for alternate calibrations with a 10 percent and 15 percent REO discount can be found in the online appendix.

Panels A and C show the maximum log change in aggregate prices and the maximum log change in retail prices, which are the two metrics that are used in the calibration of $\eta$ and $\Delta \gamma$. The model fits both well as the data points cluster around the 45-degree line, and when we regress the simulated data on the real data we get coefficients of .997 and .994, respectively.\textsuperscript{21} Another metric of fit is the mean squared error used in the simulation, which is 0.081. By comparison, turning off foreclosures and re-calibrating the model gives a mean squared error of 0.096.\textsuperscript{22} The improved fit of our structural model with foreclosures furthers the case made in Section 1 that foreclosures are a large reason why areas with larger booms had more than proportionately larger busts as illustrated in the data in Figure 3.

The remaining panels compare the model to the data on four metrics that are not used in the calibrations to provide out of sample evaluations of fit. The model performs well for the maximum change in log REO sales relative to the non-REO

\textsuperscript{21}A version of the model with a permanent shock to $a$ but not a temporary shock to $\gamma$ does a good job explaining price declines and can explain differences in the decline in retail sales across CBSAs but not the level of the decline in CBSAs. All CBSAs fall short on the decline in sales by roughly the same amount. The shock to $\gamma$ that is the same across CBSAs improves the model fit on this dimension.

\textsuperscript{22}In the no foreclosures calibration, $\eta = 1.11$ so that prices decline permanently in the bust more than they increase in the boom. If we were to restrict $\eta \in [0,1]$ when the model is re-calibrated with no foreclosures, the improvement in fit from adding foreclosures would be even greater.
Figure 7: Cross-MSA Simulations vs. Data

A: Max Change in Price Index
B: Max Change in Retail Price Index
C: Max Change in Non-REO Sales
D: Max Change in REO Sales Relative to Non-REO Peak
E: Max REO Share
F: Total Foreclosures

Note: Scatter plots of data vs simulation results for 97 CBSAs in regression analysis. The red X represents the national simulation and each black dot is a CBSA. The 45-degree line illustrates a perfect match between the model and the data. The variable being plotted shown in each plot’s title. The data is fully described in the online appendix. The calibration methodology, which fits the model to the variables in panels A and C only, is described in text.
peak, a normalized measure of REO sales, and the maximum REO share. In both cases, the points cluster around the 45-degree line and we cannot reject a coefficient of one when regressing the simulated data on actual data. The model does slightly worse for the fraction of the housing stock that is foreclosed upon. This is the case because the model underpredicts the number of foreclosures in a few CBSAs with massive price declines, which may due to an unmodeled feedback between large price declines and income shocks (Mian and Sufi, 2014). This interpretation is bolstered by the fact that the maximum change in prices is slightly under-predicted in these CBSAs. Consequently, in these places our model may under-estimate the role of foreclosures. With the exception of these few CBSAs, the model does a good job replicating the fraction of the housing stock that was foreclosed upon.

The model does not do as well explaining the maximum decline in non-REO prices. It consistently under-estimates the size of the retail price decline in the hardest-hit CBSAs, and one can reject a coefficient of one when we regress the simulated data on the empirical data. This is the case for two reasons. First, the CBSAs where the model falls short on the change in the retail price index are the same CBSAs were the simulation falls short on the number of foreclosures. Accounting for the feedback between income shocks and price declines may improve the model fit.

Second, and more importantly, there are econometric reasons why repeat sales indices such as the CoreLogic non-distressed index we use may overstate the decline in retail prices in the hardest-hit metropolitan areas. Repeat sales indices assume that the houses that sell are at all times representative of the overall housing stock. However, in the hardest-hit metropolitan areas where as much as 75 percent of sales were REOs, it is likely that the homes that were selling as retail sales were of unusually low quality. This would bias the non-REO price index downward in the hardest-hit CBSAs. By contrast, our model assumes that the econometrician has fully controlled for quality and would predict a smaller decline in non-REO prices. While the degree of bias in the price indices cannot be quantified, there is evidence that bias is present because the gap between the retail and aggregate repeat-sales price indices in the hardest-hit CBSAs is implausibly small. For instance, Figure 2 shows that in CBSAs like Las Vegas, REOs made up the vast majority of sales in the depth of the bust.

23 An alternative is a hedonic price index, which assumes that observables characteristics are valued the same over time. A hedonic index will also be biased downward if houses that are worse on unobservable characteristics are more likely to sell in the bust.
bust but that aggregate price indices only fell by a few percent more than non-retail price indices. Given the fraction of REO sales in these CBSAs and a 12.5 percent REO discount, the aggregate price index should fall at least 8 to 9 percent more than the retail price index purely due to compositional effects. However, the gap between the aggregate and retail price indices in the CoreLogic data is much narrower even though estimates of REO discounts using the same repeat-sales methodology are consistently in the 10 to 20 percent range. This discrepancy is consistent with quality bias in the bust, suggesting that the model may not be doing such a poor job in explaining retail price declines after all.

A final out of sample test evaluates an ancillary prediction of the model: that homeowners who purchase foreclosures have higher match quality than homeowners who do not. In our model, this is difficult to test because moving shocks are random, but one could extend the model so that match quality follows a Markov process and homeowners decide to sell when it falls below an endogenous threshold, as in Ngai and Sheedy (2014). Such a model would imply that homeowners who purchase foreclosures have a lower moving hazard due to better match quality. To test this prediction, the online appendix pools microdata from 29 of the 35 largest MSAs for homes purchased between 2005 and 2012 and plots moving hazards separately for REO and retail sales. Although REOs are more likely to be resold in the first year, likely due to “flippers” who purchase and renovate them, after a year the resale hazard for buyers of REOs is substantially lower than for buyers of retail sales, which is strongly suggestive of higher match quality for foreclosures.

5.4 Model Dynamics

Figure 8 illustrates the model dynamics for Chicago, which is roughly the median CBSA in terms of boom size. The figure shows the impulse response to the shocks, which reduce $a$ permanently and $\gamma$ temporarily at time zero.

As shown in panel A, at $t = 0$ prices fall considerably for both REO and retail and gradually return to steady state, particularly after five years when foreclosures taper off (panel E). The aggregate price index dips more than the retail-only price index as REO sales make up a larger fraction of sales due to the compositional effect. Panel C shows that the REO discount rises on the impact of the shocks and gradually tapers off. Panel B shows that sales fall on the impact of a shock and continue to
Note: This figure shows the results of the calibrated model using an LTV distribution corresponding to Chicago, which is approximately the median CBSA in terms of boom size. Panels A and B show the average price and sales by type, with pre-downturn price and volume normalized to 1. Panel C shows the REO discount, share of vacancies, and share of volume. Panel D shows the probability of sale conditional on a match and the unconditional probability of sale for each type with the pre-downturn probability normalized to 1. Panel E shows the annualized fraction of the owner occupied housing stock that is foreclosed upon at each point in time. Panel F shows the mass of each type of agent in the market.
fall substantially as REOs become prevalent in the market. This is the case because REOs sell more quickly, as shown by the gap between the REO share of volume and vacancies in panel C, and because the presence of REOs depletes the stock of buyers quite rapidly, as shown in panel F. The accumulation of locked-in homeowners shown in panel F also reduces sales. Finally, panel D shows how the probability of sale and probability of sale conditional on matching evolve. The gap that opens between REO and retail probabilities of sale conditional on a match reflects the choosy buyer effect. Although small in the median CBSA, this effect is more pronounced in the hardest hit CBSAs such as Las Vegas.

These dynamics reflect some important features of the data in Figures 1 and 2. For instance, REOs partly substitute for retail sales. REOs also sell faster and thus constitute a larger share of vacancies than of volume. There is also a substantial amount of conversion of previously owner-occupied homes to rentals (panel F) to accommodate rental demand by foreclosed-upon homeowners, which is consistent with both anecdotal evidence and with the fact that a substantial number of investors purchased and converted REOs to rentals. This weakens the market tightness effect relative to model without rental conversion. Additionally, the widening foreclosure discount is consistent with evidence from Campbell et al. (2011). The largest difference between the model and the data is that the model does not feature price momentum, so prices fall immediately on the impact of the shocks rather than gradually.

5.5 How Much Do Foreclosures Exacerbate the Downturn?

The calibrated model allows us to consider what the bust would have looked like in each CBSA in the absence of foreclosure. To do so, we assume that underwater homeowners who experience an income shock sell their house as a normal seller rather than defaulting and then expose each CBSA to the same set of shocks as our calibrated model with default.

Figure 9 shows the maximum log change in aggregate and retail prices for the simulation with default on the x axis and without default on the y axis. The gap between each point and the 45-degree line shows how much foreclosures exacerbated the downturn in each CBSA. Sales are not shown because there is almost no variation across CBSAs in sales without foreclosures.

Panel A shows the aggregate price index. The red X, which represents a simulation
Figure 9: Cross-MSA Simulations: Foreclosure vs. No Foreclosure

Note: This figure compares the simulate maximum log change in aggregate prices and retail prices with and without foreclosures using the same calibration, providing a counterfactual that shows the extent to which foreclosures exacerbate the downturn in the model. Each black dot is a CBSA, and the red X represents the model calibrated to the aggregate national data.

of the nation as a whole, shows that aggregate prices fall by 0.396 log points with default and 0.247 log points without. In other words, default exacerbates the decline in aggregate price indices by about 60 percent, or equivalently default accounts for about 38 percent of the decline in prices nationally. This figure is, however, larger in metropolitan areas that had larger price declines and consequently more foreclosures. For instance, the simulation of Las Vegas reveals that the price decline is 106 percent larger with foreclosures than without foreclosures.

The aggregate price index, however, includes a substantial compositional effect and is thus not the best measure of the general equilibrium effect of foreclosures on non-foreclosed homes. The best metric for this is the effect of introducing foreclosures on the price that a retail seller would get if they wanted to sell at the bottom of the market. Panel B shows this effect graphically. For the nation as a whole, foreclosures exacerbate the decline in retail prices by 24 percent or equivalently account for
roughly 20 percent of the decline in prices nationally. Again, this figure is higher in metropolitan areas that were more affected by foreclosures. Our simulation of Las Vegas shows that foreclosures exacerbate the decline in retail prices by 33 percent.

These effects are far larger than the spillover effects of foreclosures estimated by microeconometric studies that cannot account for the general equilibrium effects of foreclosure at the search market level (e.g. Campbell et al., 2011; Gerardi et al., 2012; Anenberg and Kung, 2014). An exception is Mian et al. (2014), who find larger effects using more macro variation arising from differences in foreclosure policies at state lines, which is consistent with our finding of a substantial search-market-level effect.

6 Foreclosure Policy

The substantial overshooting of prices created by introducing default implies that anti-foreclosure policy may have a significant effect on the housing market. The model can be used to quantitatively assess such policies.

A recent literature has emerged in the wake of the Great Recession that focuses on how declines in housing prices can have spillovers that dramatically affect the real economy. For instance, Iacoviello (2005) argues house price declines can have pecuniary externalities because a collapse in home prices destroys wealth in the form of home equity and can impede borrowing by households and firms, creating a financial accelerator effect similar to Kiyotaki and Moore (1997). These effects can be particularly strong in a liquidity trap. Additionally, banks may be forced to realize substantial losses on foreclosed properties, which impacts their balance sheets through the well-documented net-worth channel in financial intermediation, potentially leading to problems in the interbank repo market, cash hoarding by banks, and a freezing up of credit. Finally, there is ample evidence that debt overhang caused by a combination of high borrowing and a decline in housing prices can reduce residential investment, employment, and consumption if borrowers have a higher marginal propensity to consume (Mian and Sufi, 2011; Mian and Sufi, 2014). Given these

24 These studies compare homes with foreclosures next door to homes with foreclosures a few blocks away, so the effect of foreclosures at the search market level that we focus on is not identified. An exception is Anenberg and Kung (2014), who look at the timing of price changes of nearby homes around the time of REO listings, but they cannot account for the reduced demand at the market level due to foreclosed-upon homeowners being locked out of the market.
spillovers, understanding how foreclosure policy affects market equilibrium is crucial to understanding the efficacy of various policies despite the fact that in our model of a housing market in isolation the effects of foreclosures on aggregate welfare are relatively modest.

Policies such as prudential caps on LTVs, principal reductions such as HAMP, and monthly payment reductions such as HARP can be evaluated using the model. Doing so, however, requires estimates of the effects of policy on default rates, and there is currently no consensus on the sizes of these effects. Consequently, we focus on an analysis of a simple policy that our model is particularly well-suited to analyze: slowing down the pace of foreclosures. Such a policy has been widely discussed because it can be implemented quickly and at a low cost relative to other interventions. It has also been hotly debated since Mitt Romney proposed removing barriers that limited the pace of foreclosures in the 2012 Presidential campaign. Perhaps surprisingly, we find that slowing down foreclosures can be counterproductive.

To incorporate sluggish foreclosure into our model, we assume that when a homeowner defaults the bank begins foreclosure proceedings but that only \( \frac{1}{\phi} \) foreclosures can be processed by the system each week.\(^{25}\) The remaining loans are stuck in the foreclosure process. If prices rise and the homeowner no longer has negative equity while they are still in the foreclosure process, the foreclosure “cures” and the homeowner lists their house as a normal seller but subsequently become a renter who returns to owner-occupancy with probability \( \sigma \) because of the liquidity shock they experienced.

Figure 10 shows the aggregate and retail price indices and foreclosure starts and completions for the three values of \( \phi \): \( \phi = 0 \), which corresponds to the no-backlog case presented above, \( \phi = 6,000 \), and \( \phi = 9,000 \), which correspond to a maximum annual pace of foreclosure of 0.9 and 0.6 percent of the housing stock per year, respectively.

As can be seen in panel A, as \( \phi \) rises, the aggregate prices fall by slightly less, but retail prices by more initially and take more time to recover. The “improvement” visible in the aggregate price index thus occurs solely due to a diminished compositional effects and is entirely superficial. Retail prices, which are the correct measure

\(^{25}\)Formally, we assume that if \( f(t) \) homes are in the foreclosure process pending approval only \( \frac{f(t)}{\min\{f(t), \phi \}} \) can be processed in a given period. We choose this function as a smooth approximation to \( \min\{f(t), \frac{1}{\phi} \} \), which processes up to \( \frac{1}{\phi} \) foreclosures each period. Such an approximation is necessary for the numerical implementation.
Note: The figure shows the effect of prices and foreclosure start and completion rates for three different backlogs in a model calibrated to match Chicago, the median CBSA in the size of the price run-up. The model is calibrated to the national calibration developed in section 5. The pre-downturn price level is normalized to one. One in \( \phi \) houses can be foreclosed upon each week. \( \phi = 6,000 \) corresponds to 0.9 percent of the housing stock being foreclosed upon per year and \( \phi = 9,000 \) corresponds to 0.6 percent. The figure shows that slowing down foreclosures extends the downturn and makes prices remain low for longer. Although the overall price index rises, this is due only to a compositional effect. Indeed, the retail price index falls.

of home value for exiting homeowners because they determine home equity, fall more when foreclosures are slowed down. As can be seen in panel B, this leads to an increase in foreclosure starts and an increase in total foreclosures since prices rise and underwater homes cure slowly.

To understand why slowing down foreclosures may be counterproductive, recall that according to equation (10), price is equal to a markup over the outside option of remaining a seller next period, and according to equation (6), this outside option is the discounted value of being a seller the following period plus the probability of sale times the expected surplus form matching. Slowing down foreclosures has two effects. In the short run it reduces the strength of the market tightness term, which raises probability of sale in the short run and increases the value of being a seller.
However, it also prolongs the crisis due to the backlog slowly being released as can be seen in panel B, which reduces the value of being a seller today through the seller’s outside option of waiting to sell in the medium-term. In our simulations, this second effect dominates.

Our policy experiment, however, does not account for two countervailing forces. First, banks may respond to increased foreclosure timelines by foreclosing less and modifying loans more frequently. Second, limiting the number of foreclosures initially may improve bank balance sheets and lead to improved credit conditions today, stimulating housing demand and causing prices to rise. Our simulations show that these countervailing forces must be relatively strong in order for slowing down foreclosures to be a productive policy.

7 Conclusion

This paper argues that foreclosures play an important role in exacerbating housing downturns due to their general equilibrium effects. We develop a quantitative search model of the housing market in which banks sell foreclosed homes at a discount due to high holding costs and homeowners who are foreclosed upon cannot immediately purchase another home. Foreclosures affect market behavior by reducing the number of buyers relative to sellers and increasing the probability that a buyer meets a REO seller. Sellers, and particularly REO sellers, become highly motivated to sell and lower their prices, while buyers become more choosy and likely to walk away from a deal. Both effects create downward pressure on retail prices, which can lead to price-default spirals because price declines lead to additional default.

A calibration of the full model to cross-market data is successful in matching the cross section of price and volume declines during the crisis. The model is also able to replicate observed cross-sectional patterns in REO share and the foreclosure rate. A counterfactual analysis reveals that relative to a world without foreclosure, foreclosures exacerbated the downturn in nation-wide aggregate price indices in the recent downturn by 60 percent and in retail price indices by 24 percent. A quantitative policy experiment demonstrates that delaying foreclosures may be counterproductive. Additionally, as more experimental studies of the effectiveness of foreclosure policies at the micro level are published, the model can be used to assess the aggregate market impact of such policies.
References


A Appendix (For Online Publication)

A.1 Laws of Motion

We first describe the laws of motion for the pre-crisis housing market model with no defaults:

\[
\begin{align*}
  l_0(t + 1) &= (1 - \gamma) l_0(t) + v_b(t) q_b(\mu(t)) (1 - F_t(h_n(t))) \\
  v_v(t + 1) &= \gamma l_0(t) + v_n(t) [1 - q_s(\mu(t)) (1 - F_t(h_n(t)))] + v_a(t) \\
  v_b(t + 1) &= \gamma l_0(t) + v_h(t) [1 - q_h(\mu(t)) (1 - F_t(h_n(t)))] \\
  v_n(t) + v_a(t) &= v_v(t) \\
  v_{rs} + v_a(t) &= \zeta v_b(t)
\end{align*}
\]

Here \( v_v(t) \) reflects the stock of vacant houses which are either listed or rented out in any given period, \( v_a(t) \) is the stock of those houses which are rented out, and \( v_{rs} \) is the stock of dedicated rental units which cannot be sold as owner-occupied homes. The first equation describes the law of motion for homeowners, the second describes the law of motion of vacant houses, and the third describes the law of motion for buyers. The next is a simple adding up constraint and the final equation reflects market clearing in the rental market.

For the extended crisis model with default, the laws of motion are more complicated. They are as follows:

\[
\begin{align*}
  l_0(t + 1) &= (1 - \gamma) l_0(t) + v_b(t) q_b(\mu(t)) \sum r_m(t) (1 - F_t(h_m(t))) \\
  &\quad + w(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \\
  l_1(t + 1) &= (1 - \gamma) l_1(t) \\
  w(t + 1) &= (1 - \gamma_t) w(t) + (\gamma - \gamma_t) l_1(t) (1 - G(V_n(t))) - w(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \\
  f(t + 1) &= \gamma_t l_1(t) (1 - G(V_n(t))) + \gamma_t w(t) \\
  &\quad + f(t) \left( 1 - \frac{1}{\phi_f(t) + 1} \right) - f(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \\
  v_v(t + 1) &= \gamma l_0(t) + \gamma l_1(t) G(V_n(t)) + (w(t) + f(t)) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \\
  &\quad + v_n [1 - q_s(\mu(t)) (1 - F_t(h_n(t)))] + v_a(t) \\
  v_d(t + 1) &= \frac{f(t)}{\phi_f(t) + 1} + v_d [1 - q_s(\mu(t)) (1 - F_t(h_d(t)))]
\end{align*}
\]
\[ v_{b}(t+1) = \gamma l_0(t) + \gamma l_1(t) G(V_n(t)) + 
+ v_r(t)\sigma + v_b(t) \left[ 1 - q_b(\mu(t)) \sum r_m(t) (1 - F_l(h_m(t))) \right] \]

\[ v_{f}(t+1) = \gamma l_0(t) + \gamma l_1 G(V_n(t)) + w(t) \frac{G(V_n(t)) - G(V_n(t-1))}{1 - G(V_n(t-1))} \]

\[ + f(t) \frac{G(V_n(t)) - G(V_n(t-1))}{\phi f(t) + 1} + (1 - \sigma)v_f(t) \]

\[ v_n(t) + v_a(t) = v_v(t) \]

\[ v_a(t) + v_{rs} = \zeta [v_b(t) + v_f(t)] \]

The first two equations describe the laws of motions for homeowners. The third is the law of motion for homeowners who would like to move but are locked into their current house due to being underwater. The next provides the law of motion for the stock of foreclosed houses. The next four describe the laws of motion for vacant houses, distressed sellers, buyers, and individuals who have defaulted respectively. The next is the adding up constraint for vacant houses just as before and the final is the market clearing condition for the rental market.

### A.2 Full Dynamic Systems

For the convenience of the reader, we collect all the equations for the baseline pre-crisis model and the crisis model with default here.

The baseline pre-crisis model is:

\[ l_0(t+1) = (1 - \gamma) l_0(t) + v_b(t) q_b(\mu(t)) (1 - F_l(h_n(t))) \]

\[ v_v(t+1) = \gamma l_0(t) + v_n(t) \left[ 1 - q_s(\mu(t)) (1 - F_l(h_n(t))) \right] + v_a(t) \]

\[ v_b(t+1) = \gamma l_0(t) + v_b(t) \left[ 1 - q_b(\mu(t)) (1 - F_l(h_n(t))) \right] \]

\[ v_n(t) + v_a(t) = v_v(t) \]

\[ v_{rs} + v_a(t) = \zeta v_b(t) \]

\[ V_n(t) = m_n + \beta E_t \left( V_n(t+1) + q_s(\mu(t)) (1 - F_l(h_n(t))) \right) \frac{\theta E_t [h - h_n(t) | h \geq h_n(t)]}{1 - \beta (1 - \gamma)} \]

\[ V_d(t) = m_d + \beta E_t V_d(t+1) + q_s(\mu(t)) (1 - F_l(h_n(t))) \frac{\theta E_t [h - h_d(t) | h \geq h_d(t)]}{1 - \beta (1 - \gamma)} \]

\[ B(t) = -r(t) + \beta E_t B(t+1) \]

\[ + q_b(\mu(t)) (1 - F_l(h_n(t))) \frac{(1 - \theta) E_t [h - h_n(t) | h \geq h_n(t)]}{1 - \beta (1 - \gamma)} \]

\[ \tilde{V}_n(t) = r(t) + \beta E_t \tilde{V}_n(t+1) \]

\[ \tilde{V}_n(t) = V_n(t) \]
\[
\begin{align*}
&\frac{h_n(t)}{1 - \beta (1 - \gamma)} + \beta E_t \Gamma (t + 1) = -r(t) + \beta E_t B(t + 1) + m_n + \beta E_t V_n(t + 1) \\
&\frac{h_d(t)}{1 - \beta (1 - \gamma)} + \beta E_t \Gamma (t + 1) = -r(t) + \beta E_t B(t + 1) + m_n + \beta E_t V_d(t + 1) \\
&\Gamma(t) = \gamma(V_n(t) + B(t)) + \beta (1 - \gamma) E_t \Gamma(t + 1) \\
p_{n,h}(t) = \frac{\theta (h - h_n(t))}{1 - \beta (1 - \gamma)} + m_n + \beta E_t V_n(t + 1) \\
p_{d,h}(t) = \frac{\theta (h - h_n(t))}{1 - \beta (1 - \gamma)} + m_d + \beta E_t V_d(t + 1) \\
a(t) = \rho a(t - 1) + (1 - \rho) \bar{a} + \varepsilon(t), \ \varepsilon(t) \sim N(0, \sigma^2_\varepsilon)
\end{align*}
\]

Where for \( m = \{d, n\} \),

\[
1 - F_t(h_m(t)) = \exp[-\lambda(h_m(t) - a(t))]
\]

\[
E[h - h_n(t)|h \geq h_n(t)] = \frac{1}{\lambda}.
\]

Only three conditions are not described in the main text. The first two are of the form

\[
\frac{h_m(t)}{1 - \beta (1 - \gamma)} + \beta E_t \Gamma (t + 1) = -r(t) + \beta E_t B(t + 1) + m + \beta E_t V_m(t + 1), \ m = \{d, n\}.
\]

These are rewritten versions of

\[
V_{h_m(t)}(t) = m - r(t) + \beta \left(E_t B(t + 1) + E_t V_m(t + 1)\right), \ m = \{d, n\}
\]

that define the cutoffs for normal and REO sellers. The third is \( \Gamma(t) \), which is the continuation value of occupying a home written recursively. The model is log-linearized around a steady state in which there are no shocks and all endogenous variables are constant.

The extended model with default in Section 4 is:\textsuperscript{26}

\textsuperscript{26}These laws of motion are for the model with foreclosure backlogs, hence an additional law of motion for \( f(t) \) as a stock that drains at rate \( \frac{f(t)}{\sigma_f(t)+1} \). See footnote 25 for details.
\[ V_n(t) = m_n + \beta E_t \tilde{V}_n(t + 1) + q_s(\mu(t))(1 - F_t(h_n(t))) \frac{\theta E_t [h - h_n(t) | h \geq h_n(t)]}{1 - \beta(1 - \gamma)} \]
\[ V_d(t) = m_d + \beta E_t V_d(t + 1) + q_s(\mu(t))(1 - F_t(h_n(t))) \frac{\theta E_t [h - h_d(t) | h \geq h_d(t)]}{1 - \beta(1 - \gamma)} \]
\[ B(t) = -r(t) + \beta E_t B(t + 1) \]
\[ + \sum_{j=n,d} q_b(\mu(t)) r_j(t) (1 - F_t(h_j(t))) \frac{(1 - \theta) E_t [h - h_j(t) | h \geq h_j(t)]}{1 - \beta(1 - \gamma)} \]
\[ \Gamma(t) = \gamma(t)(V_n(t) + B(t)) + \beta(1 - \gamma(t)) E_t \Gamma(t + 1) \]

\[ \tilde{V}_n(t) = r(t) + \beta E_t \tilde{V}_n(t + 1) \]
\[ V_n(t) = \tilde{V}_n(t) \]
\[ l_0(t + 1) = (1 - \gamma(t)) l_0(t) + v_b(t) q_b(\mu(t)) \sum r_m(t) (1 - F_t(h_m(t))) \]
\[ + w(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \]
\[ l_1(t + 1) = (1 - \gamma(t)) l_1(t) \]
\[ w(t + 1) = (1 - \gamma_f) w(t) + (\gamma(t) - \gamma_f) l_1(t) (1 - G(V_n(t))) \]
\[ - w(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \]
\[ f(t + 1) = \gamma l_1(t) (1 - G(V_n(t))) + \gamma_f w(t) \]
\[ + f(t) \left( 1 - \frac{1}{\phi f(t) + 1} \right) - f(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \]
\[ v_v(t + 1) = \gamma(t) l_0(t) + \gamma(t) l_1(t) G(V_n(t)) + (w(t) + f(t)) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \]
\[ + v_n [1 - q_s(\mu(t))(1 - F_t(h_n(t)))] + v_a(t) \]
\[ v_d(t + 1) = \frac{f(t)}{\phi f(t) + 1} + v_d [1 - q_s(\mu(t))(1 - F_t(h_d(t)))] \]
\[ v_b(t + 1) = \gamma(t) l_0(t) + \gamma(t) l_1(t) G(V_n(t)) + \]
\[ + v_r(t) \sigma + v_b(t) \left[ 1 - q_b(\mu(t)) \sum r_m(t) (1 - F_t(h_m(t))) \right] \]
\[ v_f(t + 1) = \gamma l_0(t) + \gamma l_1 G(V_n(t)) + w(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} \]
\[ + f(t) \frac{G(V_n(t)) - G(V_n(t - 1))}{1 - G(V_n(t - 1))} + \frac{f(t)}{\phi f(t) + 1} + (1 - \sigma) v_f(t) \]
\[ v_n(t) + v_a(t) = v_v(t) \]
\[ v_a(t) + v_{rs} = \zeta [v_b(t) + v_f(t)] \]
\[
\begin{align*}
\frac{h_n(t)}{1 - \beta (1 - \gamma(t))} + \beta E_t \Gamma(t + 1) &= -r(t) + \beta E_t B(t + 1) + m_n + \beta E_t V_n(t + 1) \\
\frac{h_d(t)}{1 - \beta (1 - \gamma(t))} + \beta E_t \Gamma(t + 1) &= -r(t) + \beta E_t B(t + 1) + m_n + \beta E_t V_d(t + 1) \\
\gamma(t) &= \rho \gamma(t - 1) + (1 - \rho) \bar{\gamma} + \Delta \gamma \\
a(t) &= \bar{a} + \Delta a
\end{align*}
\]

where $\Delta \gamma$ and $\Delta a$ are the temporary shock to $\gamma$ and the permanent shock to $a$, respectively. $\Delta a$ is chosen so that the permanent price decline due to the shock is $\Delta p$ defined in the main text.

To capture the rich nonlinear dynamics implied by default, further stochastic shocks to $a$ are shut down and the dynamic transition path between the two steady-states is solved for using a Newton algorithm. In each steady state all variables are constant given the long-run values for $a$. Both the pre-crisis and crisis models are solved using the Dynare software package.

### A.3 Alternate REO Discounts

The main text uses a 12.5 percent REO discount, but this number cannot be verified with certainty due to unobserved quality. This section shows results when the model is re-calibrated to a 10 percent discount and a 15 percent discount. Qualitatively, the results are very similar, so rather than focusing on minute changes in the model’s dynamics we present the model’s fit as in Figure 7 and the no foreclosure counterfactual as in Figure 9. The fit to the pre-downturn moments and parameters do not change dramatically and are available upon request.

### A.4 10 Percent Discount

Figure 11 shows the model fit for a 10 percent discount. Of note is that the model does worse at predicting the change in the aggregate price index for CBSAs that experienced a large price decline. This is the case both because the compositional effect is weakened with a 10 percent discount and because the increase in the foreclosure discount is smaller. Consequently, the model fit is not nearly as good with a 10 percent foreclosure discount as it is in the 12.5 percent discount case. The model fit is also somewhat worse because with a 10 percent discount the minimum mean squared error solution sacrifices fit on the decline in price to improve fit on the decline retail sales volume.

In the counterfactuals presented in Figure 12, foreclosures exacerbate the downturn in prices by 73 percent and in retail prices by 43 percent. The effects are larger than in the main text because the permanent price decline without foreclosures is smaller as the solution sacrifices fit on price to improve fit on sales volume. Conse-
Figure 11: Cross-MSA Simulations vs. Data: 10 Percent Discount

Note: Scatter plots of data vs simulation results for 97 CBSAs in regression analysis. The red X represents the national simulation and each black dot is a CBSA. The 45-degree line illustrates a perfect match between the model and the data. The variable being plotted shown in each plot’s title. The data is fully described in the online appendix. The calibration methodology is described in text.
Figure 12: Foreclosure vs. No Foreclosure, 10 Percent Discount

Note: This figure compares the simulate maximum log change in aggregate prices and retail prices with and without foreclosures using the same calibration, providing a counterfactual that shows the extent to which foreclosures exacerbate the downturn in the model. Each black dot is a CBSA, and the red X represents the model calibrated to the aggregate national data. Consequently, the increase in the price decline caused by adding foreclosures is proportionally larger even though it is smaller in absolute terms.

A.5 15 Percent Discount

Figure 11 shows the model fit for the 15 percent discount. With a 15 percent discount, the model does a good job fitting both the maximum change in the price index and non-REO sales, but it falls short in explaining the maximum change in the retail price index in the hardest-hit MSAs. It does well on the other out-of-sample tests as well. Overall, the 15 percent discount looks very similar to the 12.5 percent discount in the main text.

In the counterfactuals presented in Figure 12, the effects of foreclosures are a bit stronger for the national average but much stronger in the hardest-hit MSAs. For the nation as a whole, foreclosures exacerbate the downturn in aggregate prices by 60 percent and in retail prices by 25 percent, compared with 60 percent and 24 percent.
Figure 13: Cross-MSA Simulations vs. Data: 15 Percent Discount

Note: Scatter plots of data vs simulation results for 97 CBSAs in regression analysis. The red X represents the national simulation and each black dot is a CBSA. The 45-degree line illustrates a perfect match between the model and the data. The variable being plotted shown in each plot’s title. The data is fully described in the online appendix. The calibration methodology is described in text.
Figure 14: Foreclosure vs. No Foreclosure, 15 Percent Discount

Note: This figure compares the simulate maximum log change in aggregate prices and retail prices with and without foreclosures using the same calibration, providing a counterfactual that shows the extent to which foreclosures exacerbate the downturn in the model. Each black dot is a CBSA, and the red X represents the model calibrated to the aggregate national data.

in the main text. However, in a hard-hit city like Las Vegas, foreclosures exacerbate the downturn in aggregate prices by 161 percent and in retail prices by 56 percent, relative to 106 percent and 33 percent in the main text.

In sum, the 10 percent and 15 percent discounts show that the results are robust to the foreclosure discount, although the model fit is weakened with the smaller discount and the extent to which foreclosures exacerbate the downturn in the hardest-hit metropolitan areas is somewhat sensitive to the discount.

A.6 Data Sources and Calculations

National and Cross-CBSA Data

The main data source is proprietary data from CoreLogic, which we supplement with data from the U.S. Census, Saiz (2010), and the Wharton Land-Use Regulation Survey.

CoreLogic provides us with a monthly data set for the nation and the 100 largest
CBSAs\textsuperscript{27} for 2000-2013 compiled from public records and mortgage data. CoreLogic estimates that its data covers 85 percent of the U.S. Our data set includes:

- The CoreLogic home price index and non-distressed home price index estimated from public records. In cases where public records do not include price (misreported observations or states where the price is not disclosed), this data is supplemented with data on individual mortgages that includes purchase prices. We refer to these as the aggregate and retail price indices. The CoreLogic non-distressed price index drops REO sales and short sales from the database and re-estimates the price index using the same methodology.

- The number of pre-foreclosure filings and completed foreclosure auctions esti-

\textsuperscript{27}By CBSA code and name, they are: 10420 Akron, OH; 10580 Albany-Schenectady-Troy, NY; 10740 Albuquerque, NM; 10900 Allentown-Bethlehem-Easton, PA-NJ; 12060 Atlanta-Sandy Springs-Marietta, GA; 12420 Austin-Round Rock-San Marcos, TX; 12540 Bakersfield-Delano, CA; 12580 Baltimore-Towson, MD; 12940 Baton Rouge, LA; 13646 Bethesda-Rockville-Frederick, MD; 13820 Birmingham-Hoover, AL; 14860 Bridgeport-Stamford-Norwalk, CT; 15380 Buffalo-Niagara Falls, NY; 15764 Cambridge-Newton-Framingham, MA; 15804 Camden, NJ; 16700 Charleston-North Charleston-Summerville, SC; 16740 Charlotte-Gaston-Rock Hill, NC-SC; 16974 Chicago-Joliet-Naperville, IL; 17140 Cincinnati-Middletown, OH-KY-IN; 17460 Cleveland-Elyria-Mentor, OH; 17820 Colorado Springs, CO; 17900 Columbia, SC; 18140 Columbus, OH; 19124 Dallas-Plano-Irving, TX; 19380 Dayton, OH; 19740 Denver-Aurora-Broomfield, CO; 19904 Detroit-Livonia-Birmingham, MI; 20764 Edison-New Brunswick, NJ; 21340 El Paso, TX; 22744 Fort Lauderdale-Pompano-Beach-Deerfield Beach, FL; 23104 Fort Worth-Arlington, TX; 23420 Fresno, CA; 23844 Gary, IN; 24340 Grand Rapids-Wyoming, MI; 24660 Greensboro-High Point, NC; 24860 Greenville-Mauldin-Easley, SC; 25540 Hartford-West Hartford-East Hartford, CT; 26180 Honolulu, HI; 26420 Houston-Sugar Land-Baytown, TX; 26900 Indianapolis-Carmel, IN; 27260 Jacksonville, FL; 28140 Kansas City, MO-KS; 28940 Kentucky-Louisville-New Albany, KY-IN; 29404 Lake County-Kenosha County, IL-WI; 29820 Las Vegas-Paradise, NV; 30780 Little Rock-North Little Rock-Conway, AR; 31084 Los Angeles-Long Beach-Glendale, CA; 31140 Louisville-Jefferson County, KY-IN; 32580 McAllen-Edinburg-Mission, TX; 32820 Memphis, TN-MS-AR; 33124 Miami-Miami Beach-Kendall, FL; 33340 Milwaukee-Waukesha-West Allis, WI; 33460 Minneapolis-St. Paul-Bloomington, MN-WI; 34980 Nashville-Davidson-Murfreesboro-Franklin, TN; 35004 Nassau-Suffolk, NY; 35084 Newark-Union, NJ-PA; 35300 New Haven-Milford, CT; 35380 New Orleans-Metairie-Kenner, LA; 35644 New York-White Plains-Wayne, NY-NJ; 35840 North Port-Bradenton-Sarasota, FL; 36084 Oakland-Fremont-Hayward, CA; 36420 Oklahoma City, OK; 36540 Omaha-Council Bluffs, NE-IA; 36740 Orlando-Kissimmee-Sanford, FL; 37100 Oxnard-Thousand Oaks-Ventura, CA; 37764 Peabody, MA; 37964 Philadelphia, PA; 38060 Phoenix-Mesa-Glendale, AZ; 38300 Pittsburgh, PA; 38900 Portland-Vancouver-Hillsboro, OR-WA; 39100 Poughkeepsie-Newburgh-Middletown, NY; 39300 Providence-New Bedford-Fall River, RI-MA; 39580 Raleigh-Cary, NC; 40060 Richmond, VA; 40140 Riverside-San Bernardino-Ontario, CA; 40380 Rochester, NY; 40900 Sacramento–Arden-Arcade–Roseville, CA; 41180 St. Louis, MO-IL; 41620 Salt Lake City, UT; 41700 San Antonio-New Braunfels, TX; 41740 San Diego-Carlsbad-San Marcos, CA; 41884 San Francisco-San Mateo-Redwood City, CA; 41940 San Jose-Sunnyvale-Santa Clara, CA; 42044 Santa Ana-Anaheim-Irvine, CA; 42644 Seattle-Bellevue- Everett, WA; 44140 Springfield, MA; 44700 Stockton, CA; 45060 Syracuse, NY; 45104 Tacoma, WA; 45300 Tampa-St. Petersburg-Clearwater, FL; 45780 Toledo, OH; 46060 Tucson, AZ; 46140 Tulsa, OK; 47260 Virginia Beach-Norfolk-Newport News, VA-NC; 47644 Warren-Troy-Farmington Hills, MI; 47894 Washington-Arlington-Alexandria, DC-VA-MD-WV; 48424 West Palm Beach-Boca Raton-Boynton Beach, FL; 48864 Wilmington, DE-MD-NJ; 49340 Worcester, MA.
• Sales counts for REOs, new houses, non-REO and non-short sale resales, and short sales estimated from public records. Because short sales are not reported separately for much of the time frame covered by the data, we combine short sales and resales into a non-REO existing home sales measure which we call retail sales. We calculate existing home sales by adding REO and retail sales. We also use this data to construct the REO share of existing home volume, which we seasonally adjust.

• Estimates of 7 quantiles of the combined loan-to-value distribution for active mortgages: under 50%, 50%-60%, 60%-70%, 70%-80%, 80%-90%, 90%-100%, 100%-110%, and over 110%. These statistics are compiled by CoreLogic using public records and CoreLogic’s valuation models.

We seasonally adjust the raw CoreLogic house price indices, foreclosure counts, sales counts, and delinquent and in-foreclosure loan shares using the Census Bureau’s X-12 ARIMA software with an additive seasonal factor. For the CBSA sales counts, auctions counts, days on the market, and REO share, we smooth the data using a 5 month moving average (2 months prior, the current month, and 2 months post) to remove spikes in the data caused by irregular reporting at the county level.

For the calibration of the loan balance distribution and initial number of mortgages with high LTV ratios, we adjust the CoreLogic data using data from the American Community Survey as tabulated by the Census. The CoreLogic data only covers all active loans, while our model corresponds to the entire owner-occupied housing stock. Consequently, we use the ACS 3-year 2005-2007 estimates of the owner-occupied housing stock and fraction of houses with a mortgage at the national and county level, which we aggregate to the CBSA level using CBSA definitions. From this data, we construct the fraction of owner-occupied housing units with a mortgage and the fraction of owner-occupied housing units with a second lien or home equity loan. We use these estimates to adjust the loan balance distribution so it represents the entire owner-occupied housing stock and in our regressions to construct the fraction of owner-occupied houses with over 80 percent LTV.

The LTV data is first available for March 2006, which roughly corresponds to the eve of the housing bust as the seasonally-adjusted national house price index reached its peak in March 2006. To approximate the size of the run-up, we average the seasonally-adjusted price index for March-May 2003 and March-May 2006 to calculate the change in log prices for 2003 to 2006.

28The 3-year ACS estimates include estimates of the housing stock and houses with a mortgage for all counties with over 20,000 residents. For a few MSAs, one or more small counties are not included in the ACS data. The bias on our constructed estimates of the fraction of owner-occupied homes with a mortgage and with a second lien or home equity loan due to these small missing counties is minimal.
We also calculate the maximum peak-to-trough log change in seasonally-adjusted aggregate and retail prices, smoothed and seasonally-adjusted non-REO volume, and the maximum REO share for each geographical area. We also construct a measure of the increase in REO sales, which is the maximum monthly number of REO sales relative to the peak of non-REO sales. We use this instead of the log change in REO sales because there were so few REO sales before the crisis. We estimate the minimum value between March 2006 and September 2013 and the maximum value between January 2002 and December 2007. We implement these restrictions so that the addition of counties to the CoreLogic data set prior to 2002 does not distort our results. We calculate the fraction of the owner-occupied housing stock that was foreclosed upon by adding up completed foreclosure auctions between March 2006 and September 2013 and dividing by the owner-occupied housing stock in 2006 as calculated from the ACS. Again, our results are not sensitive to the choice of dates.

From the 100 CBSAs we drop the Syracuse, New York CBSA which has incomplete data. We also exclude the Birmingham, Alabama CBSA from calculations on sales volume is dropped because a major county stopped reporting to CoreLogic in the middle of the downturn so the sales volume series is discontinuous. In the main text we exclude two CBSAs in the greater Detroit area—Detroit and Warren—because they experienced large price declines without a prior boom and thus create an exaggerated non-linear relationship between boom size and bust size. Below we show results including these two CBSAs, which are robust to including them. We thus have 97 CBSAs for price and 96 for sales volume.

For robustness tests, we merge data from Saiz (2010) into the CBSA data. The Saiz data includes his estimate of unusable land due to terrain, the housing supply elasticity, and the Wharton Land-Use Regulation Survey score for each CBSA. We are able to match every CBSA we have data on except for Sacramento CA and Honolulu HI. Summary statistics for the complete data set are in Table 6.

As a robustness test, Table 7 shows regression results when equation (1) is augmented to include the fraction of homes with a second mortgage in 2006, the land unavailability index, and the land use regulation index. We add the fraction of individuals with a second mortgage or home equity loan to the regression because these loans have received attention in analyses of the downturn (Mian and Sufi, 2011) and we want to make sure they are not driving the result. We use a land unavailability index and the Wharton land use regulation index from Saiz (2010) to proxy for the housing supply elasticity. The resulting regression is:

\[
Y = \beta_0 + \beta_1 \max_{03-06} \log (P) + \beta_2 [\Delta_{03-06} \log (P)]^2 + \beta_3 (Z \text{ max Share LTV > 80%}) + \beta_4 (\Delta_{03-06} \log (P) \times Z \text{ LTV > 80%}) + \beta_5 (Z \% \text{ Second Mortgage, 2006}) + \beta_6 (\Delta_{03-06} \log (P) \times Z \% \text{ Second}) + \beta_7 (Z \text{ Saiz Land Unavailability}) + \beta_8 (Z \text{ Wharton Land Use Regulation}) + \varepsilon. \tag{12}
\]
Table 6: MSA Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Unweighted Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max $\Delta \log (P)$</td>
<td>-0.326</td>
<td>0.219</td>
<td>-0.950</td>
<td>-0.028</td>
<td>97</td>
</tr>
<tr>
<td>Max $\Delta \log (P_{\text{Retail}})$</td>
<td>-0.278</td>
<td>0.195</td>
<td>-0.880</td>
<td>-0.035</td>
<td>97</td>
</tr>
<tr>
<td>Max $\Delta \log (\frac{\text{Sales}<em>{\text{Retail}}}{\text{Sales}</em>{\text{Existing}}})$</td>
<td>-1.124</td>
<td>0.295</td>
<td>-1.918</td>
<td>-0.390</td>
<td>96</td>
</tr>
<tr>
<td>$\frac{\text{Sales}<em>{\text{REO}}}{\text{Sales}</em>{\text{Existing}}}$</td>
<td>0.307</td>
<td>0.163</td>
<td>0.084</td>
<td>0.796</td>
<td>96</td>
</tr>
<tr>
<td>% Foreclosed</td>
<td>0.091</td>
<td>0.073</td>
<td>0.011</td>
<td>0.438</td>
<td>96</td>
</tr>
<tr>
<td>$\Delta \log (\text{Price})_{03-06}$</td>
<td>0.306</td>
<td>0.181</td>
<td>0.038</td>
<td>0.729</td>
<td>97</td>
</tr>
<tr>
<td>Share LTV &gt; 80%</td>
<td>0.143</td>
<td>0.076</td>
<td>0.026</td>
<td>0.328</td>
<td>97</td>
</tr>
<tr>
<td>Frac Second Mort, 06</td>
<td>0.203</td>
<td>0.053</td>
<td>0.026</td>
<td>0.290</td>
<td>97</td>
</tr>
<tr>
<td>Saiz Land Unav</td>
<td>0.280</td>
<td>0.213</td>
<td>0.009</td>
<td>0.796</td>
<td>95</td>
</tr>
<tr>
<td>Wharton Land Reg</td>
<td>0.228</td>
<td>0.711</td>
<td>-1.239</td>
<td>1.892</td>
<td>95</td>
</tr>
</tbody>
</table>

Notes: Summary statistics for variables used in regression analysis. All data is from CoreLogic. Data is for the 100 largest CBSAs not including Syracuse NY and Detroit and Warren MI. For sales, the REO share and percent foreclosed, Birmingham AL is also omitted. For the Saiz land unavailability and the Wharton land-use regulation, Sacramento CA and Honolulu HI are omitted.

The key patterns in the main text continue to be present in this table. Table 8 shows the results are robust to including the two outlier CBSAs in the greater Detroit area that are dropped from the main analysis.

**DataQuick Data**

We supplement the national and CBSA-level data from CoreLogic with microdata from DataQuick. DataQuick collects and digitizes public records from county register of deeds and assessor offices and provides a rich one-time snapshot of housing characteristics from 2013 along with a detailed transaction history of each property from 1988 to 2013 (1997 to 2013 in some CBSAs) that includes transaction prices, loans, buyer and seller names and characteristics, and seller distress. The data is cleaned to remove problematic observations and duplicates as detailed in the data appendix of Guren (2013). We limit our analysis to resale transactions (as opposed to new construction or subdivisions) of residential homes (both single- and multi-family), both of which are categorized by DataQuick.

We use the DataQuick data to conduct an analysis of whether individuals who buy foreclosed homes are more or less likely to stay in their home. To do so, we merge data from the 29 of the 35 largest CBSAs in Table 9. We consider each purchase of a home as a spell that lasts until the house is sold. Since the data is censored, we use a duration analysis. We have 11,922,102 spells of individuals who purchased between January 1 2005 and May 31, 2012.

Figure 15 nonparametrically plots the resale hazard function separately for REOs and non-REOs for properties purchased from 2005 to 2012. Although REOs have a
<table>
<thead>
<tr>
<th></th>
<th>$\Delta \log (P)$ 03–06</th>
<th>$\Delta \log (P_{\text{Retail}})$</th>
<th>$\Delta \log (\text{Sales}_{\text{Retail}})$</th>
<th>$\Delta \text{Sales}_{\text{REO}}$</th>
<th>$% \text{ Foreclosed}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \log (\text{Price})$ 03–06</td>
<td>0.586</td>
<td>0.374</td>
<td>-1.119</td>
<td>-1.168</td>
<td>0.430</td>
</tr>
<tr>
<td></td>
<td>(0.403)</td>
<td>(0.305)</td>
<td>(0.644)**</td>
<td>(0.382)**</td>
<td>(0.131)**</td>
</tr>
<tr>
<td>$\Delta \log (\text{Price})^2$ 03–06</td>
<td>-2.234</td>
<td>-1.843</td>
<td>-0.489</td>
<td>2.440</td>
<td>1.075</td>
</tr>
<tr>
<td></td>
<td>(0.490)**</td>
<td>(0.370)**</td>
<td>(0.747)</td>
<td>(0.454)**</td>
<td>(0.162)**</td>
</tr>
<tr>
<td>$Z \text{ Share LTV} &gt; 80%$</td>
<td>0.062</td>
<td>0.068</td>
<td>-0.004</td>
<td>-0.008</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.029)**</td>
<td>(0.025)**</td>
<td>(0.049)</td>
<td>(0.030)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>$\Delta \log (P) \times Z \text{ LTV} &gt; 80%$</td>
<td>-0.330</td>
<td>-0.346</td>
<td>-0.352</td>
<td>0.236</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td>(0.113)**</td>
<td>(0.112)**</td>
<td>(0.146)**</td>
<td>(0.093)**</td>
<td>(0.060)**</td>
</tr>
<tr>
<td>$Z \text{ Frac Second Mort, 06}$</td>
<td>-0.063</td>
<td>-0.044</td>
<td>-0.109</td>
<td>-0.018</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>(0.020)**</td>
<td>(0.018)**</td>
<td>(0.035)**</td>
<td>(0.025)</td>
<td>(0.009)*</td>
</tr>
<tr>
<td>$\Delta \log (P) \times Z \text{ Second}$</td>
<td>0.118</td>
<td>0.062</td>
<td>0.104</td>
<td>0.231</td>
<td>0.066</td>
</tr>
<tr>
<td></td>
<td>(0.076)</td>
<td>(0.073)</td>
<td>(0.102)</td>
<td>(0.080)**</td>
<td>(0.032)**</td>
</tr>
<tr>
<td>$Z \text{ Saiz Land Unav}$</td>
<td>-0.007</td>
<td>0.001</td>
<td>-0.008</td>
<td>-0.004</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.013)</td>
<td>(0.019)</td>
<td>(0.015)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>$Z \text{ Wharton Land Reg}$</td>
<td>-0.025</td>
<td>-0.022</td>
<td>-0.003</td>
<td>0.017</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.011)**</td>
<td>(0.010)**</td>
<td>(0.017)</td>
<td>(0.011)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.259</td>
<td>-0.196</td>
<td>-0.762</td>
<td>0.376</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>(0.066)**</td>
<td>(0.050)**</td>
<td>(0.110)**</td>
<td>(0.065)**</td>
<td>(0.021)**</td>
</tr>
<tr>
<td>$r^2$</td>
<td>0.764</td>
<td>0.777</td>
<td>0.680</td>
<td>0.562</td>
<td>0.666</td>
</tr>
<tr>
<td>$N$</td>
<td>95</td>
<td>95</td>
<td>94</td>
<td>94</td>
<td>94</td>
</tr>
</tbody>
</table>

Notes: * = 10% Significance, ** = 5% Significance *** = 1% significance. All standard errors are robust to heteroskedasticity. Each column shows estimates of (1) as in Table 1 but with no coefficients suppressed. All data is from CoreLogic and fully described in Appendix A.6. Data is for 100 largest CBSAs excluding Syracuse NY, Detroit and Warren MI which are removed so that the outliers do not exaggerate the non-linear relationship, and for columns 3-5 Birmingham AL, which has an inconsistent sales series.
Table 8: Cross MSA Regressions on the Impact of the Size of the Bubble and Its Interaction With High LTV Share: Including Outliers

<table>
<thead>
<tr>
<th></th>
<th>( \Delta \log (P) )</th>
<th>( \Delta \log (P_{\text{Retail}}) )</th>
<th>( \Delta \log (\text{Sales}_{\text{Retail}}) )</th>
<th>( \frac{\text{Sales}<em>{\text{RFO}}}{\text{Sales}</em>{\text{Existing}}} )</th>
<th>% Foreclosed</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \log (\text{Price})_{03-06} )</td>
<td>1.213</td>
<td>0.903</td>
<td>-0.262</td>
<td>-1.644</td>
<td>-0.631</td>
</tr>
<tr>
<td></td>
<td>(0.571)**</td>
<td>(0.468)*</td>
<td>(0.813)</td>
<td>(0.471)**</td>
<td>(0.194)*****</td>
</tr>
<tr>
<td>( \Delta \log (\text{Price})_{03-06}^2 )</td>
<td>-2.942</td>
<td>-2.441</td>
<td>-1.459</td>
<td>2.978</td>
<td>1.303</td>
</tr>
<tr>
<td></td>
<td>(0.683)**</td>
<td>(0.558)**</td>
<td>(0.941)</td>
<td>(0.554)**</td>
<td>(0.233)*****</td>
</tr>
<tr>
<td>( Z \text{ Share LTV &gt; 80%} )</td>
<td>0.057</td>
<td>0.063</td>
<td>-0.007</td>
<td>-0.006</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.030)**</td>
<td>(0.056)</td>
<td>(0.033)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>( \Delta \log (P) \times Z \text{ LTV &gt; 80%} )</td>
<td>-0.277</td>
<td>-0.299</td>
<td>-0.291</td>
<td>0.201</td>
<td>0.202</td>
</tr>
<tr>
<td></td>
<td>(0.139)**</td>
<td>(0.134)**</td>
<td>(0.173)*</td>
<td>(0.106)*</td>
<td>(0.067)*****</td>
</tr>
<tr>
<td>( Z \text{ Frac Second Mort, 06} )</td>
<td>-0.054</td>
<td>-0.034</td>
<td>-0.104</td>
<td>-0.021</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>(0.028)*</td>
<td>(0.024)</td>
<td>(0.044)**</td>
<td>(0.028)</td>
<td>(0.011)*</td>
</tr>
<tr>
<td>( \Delta \log (P) \times Z \text{ Second} )</td>
<td>0.079</td>
<td>0.025</td>
<td>0.073</td>
<td>0.250</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td>(0.097)</td>
<td>(0.089)</td>
<td>(0.126)</td>
<td>(0.089)**</td>
<td>(0.037)****</td>
</tr>
<tr>
<td>( Z \text{ Saiz Land Unav} )</td>
<td>-0.017</td>
<td>-0.007</td>
<td>-0.021</td>
<td>0.003</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.020)</td>
<td>(0.015)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>( Z \text{ Wharton Land Reg} )</td>
<td>-0.029</td>
<td>-0.025</td>
<td>-0.006</td>
<td>0.019</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>(0.012)**</td>
<td>(0.011)**</td>
<td>(0.018)</td>
<td>(0.012)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.363</td>
<td>-0.283</td>
<td>-0.905</td>
<td>0.456</td>
<td>0.142</td>
</tr>
<tr>
<td></td>
<td>(0.093)**</td>
<td>(0.076)**</td>
<td>(0.139)**</td>
<td>(0.080)**</td>
<td>(0.031)*****</td>
</tr>
<tr>
<td>( r^2 )</td>
<td>0.689</td>
<td>0.704</td>
<td>0.606</td>
<td>0.524</td>
<td>0.614</td>
</tr>
<tr>
<td>( N )</td>
<td>97</td>
<td>97</td>
<td>96</td>
<td>96</td>
<td>96</td>
</tr>
</tbody>
</table>

Notes: * = 10% Significance, ** = 5% Significance *** = 1% significance. All standard errors are robust to heteroskedasticity. All data is from CoreLogic and fully described in Appendix A.6. Data is for 100 largest CBSAs excluding Syracuse NY, Sacramento CA, Honolulu HI, and for columns 3-5 Birmingham AL, which has an inconsistent sales series.
## Table 9: Metro Areas in DataQuick Data

<table>
<thead>
<tr>
<th>Metro Area</th>
<th>Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Los Angeles</td>
<td>Los Angeles CA, Orange CA</td>
</tr>
<tr>
<td>Chicago</td>
<td>Cook IL, DuPage IL, Kane IL, Kendall IL, Lake IL, McHenry IL, Will IL, Lake IN, Porter IN</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>New Castle DE, Cecil MD, Burlington NJ, Camden NJ, Gloucester NJ, Salem NJ, Bucks PA, Chester PA, Delaware PA, Montgomery PA, Philadelphia PA</td>
</tr>
<tr>
<td>DC</td>
<td>District of Columbia, Montgomery MD, Prince George's MD, Alexandria VA, Arlington VA, Fairfax VA, Falls Church VA, Fauquier VA, Loudoun VA, Manassas VA, Manassas Park VA, Prince William VA, Stafford VA</td>
</tr>
<tr>
<td>Miami</td>
<td>Miami-Dade FL, Broward FL, Palm Beach FL</td>
</tr>
<tr>
<td>Boston</td>
<td>Essex MA, Middlesex MA, Norfolk MA, Plymouth MA, Suffolk MA, Rockingham NH</td>
</tr>
<tr>
<td>San Francisco</td>
<td>Alameda CA, Marin CA, San Francisco CA, San Mateo CA</td>
</tr>
<tr>
<td>Riverside</td>
<td>Riverside CA, San Bernardino CA</td>
</tr>
<tr>
<td>Phoenix</td>
<td>Maricopa AZ, Pinal AZ</td>
</tr>
<tr>
<td>Detroit</td>
<td>Macomb MI, Oakland MI, Wayne MI</td>
</tr>
<tr>
<td>Seattle</td>
<td>King WA, Pierce WA, Snohomish WA</td>
</tr>
<tr>
<td>Minneapolis</td>
<td>Anoka MN, Carver MN, Dakota MN, Hennepin MN, Ramsey MN, Scott MN, Sherburne MN, Washington MN, Wright MN</td>
</tr>
<tr>
<td>San Diego</td>
<td>San Diego CA</td>
</tr>
<tr>
<td>Tampa</td>
<td>Hillsborough FL, Pasco FL, Pinellas FL</td>
</tr>
<tr>
<td>St. Louis</td>
<td>Madison IL, Monroe IL, St Clair IL, Jefferson MO, St Charles MO, St Louis MO, St Louis City MO</td>
</tr>
<tr>
<td>Baltimore</td>
<td>Anne Arundel MD, Baltimore City MD, Baltimore MD, Howard MD, Queen Anne's MD</td>
</tr>
<tr>
<td>Denver</td>
<td>Adams CO, Arapahoe CO, Broomfield CO, Denver CO, Douglas CO, Jefferson CO</td>
</tr>
<tr>
<td>Pittsburgh</td>
<td>Allegheny PA, Beaver PA, Butler PA, Washington PA, Westmoreland PA</td>
</tr>
<tr>
<td>Charlotte</td>
<td>Iredell NC, Mecklenburg NC, Union NC</td>
</tr>
<tr>
<td>Portland</td>
<td>Clackamas OR, Multnomah OR, Washington OR, Clark WA</td>
</tr>
<tr>
<td>Orlando</td>
<td>Orange FL, Seminole FL</td>
</tr>
<tr>
<td>Sacramento</td>
<td>El Dorado CA, Placer CA, Sacramento CA</td>
</tr>
<tr>
<td>Cincinnat</td>
<td>KY, Kenton KY, Butler OH, Clermont OH, Hamilton OH, Warren OH</td>
</tr>
<tr>
<td>Cleveland</td>
<td>Cuyahoga OH, Geauga OH, Lake OH, Medina OH</td>
</tr>
<tr>
<td>Las Vegas</td>
<td>Clark NV</td>
</tr>
<tr>
<td>Columbus</td>
<td>Delaware OH, Fairfield OH, Franklin OH</td>
</tr>
<tr>
<td>San Jose</td>
<td>Santa Clara CA</td>
</tr>
</tbody>
</table>
much higher resale hazard for the first year, suggesting they are much more likely to be “flipped” by developers, after one year the resale hazard is substantially lower than for retail homes. This suggests that match quality is higher for REOs purchased during the downturn, which is consistent with our model. Indeed, if we extended our model to have an endogenous moving decision rather than a fixed per-period moving probability, one would expect to see exactly this type of relationship.

**Target Moments**

This section provides details on the target moments used in the calibration in Section 3.

We target a mean house price for a retail sales of $300,000 as an approximation to Adelino et al.’s (2012) mean house price of $298,000 for 10 MSAs. In reporting results, we normalize this initial house price to 1. The results are not sensitive to this figure, which is a normalization.
As discussed in the text, REO discounts are hard to estimate due to unobserved quality. Most estimates of REO discounts prior to the downturn were approximately 20 percent, but some estimates are closer to 10 percent (see Clauretie and Denshvary, 2009 and Campbell et al., 2011). In the main text we use 12.5 percent, which is a conservative figure that attributes a substantial amount of the discount to unobserved quality, and in the online appendix we present results for 10 percent and 15 percent discounts.

We target a time on the market for retail houses of 26 weeks as in Piazzesi and Schneider (2009). This number is a bit higher than some papers that use Multiple Listing Service Data such as Anenberg (2013) and Springer (1996), likely because of imperfect adjustment for withdrawn listings and re-listings.

We target a time on the market for REO houses of 15 weeks. Springer (1996) analyzes various forms of “seller motivation” such as relocation and financial distress using data form Texas from 1991-3. He finds that a foreclosure sales are the only motivated sellers that have significantly reduced time on the market. His estimate is that time on the market is reduced by .2135 log points or 23.7%. However, REO sales are almost never withdrawn from the market, whereas retail sales are frequently withdrawn (Anenberg, 2013). We also attempt to adjust so our number excludes extremely rundown properties that sit on the market for several years.

We target a ratio of buyer to seller time on the market is 1.117 from Genesove and Han’s (2012) analysis of National Association of Realtor surveys.

We target a standard deviation of annual log changes in prices is of .047 to match the CoreLogic national house price index adjusted for the CPI from 1976-2006.

We target a standard deviation of annual log changes of .133 to match the standard deviation of annual log changes in sales in National Association of Realtors data, which is 2.82 times the standard deviation of annual log changes in prices as measured by the CoreLogic national house price index from 1976-2006.

The median tenure for owner occupants of approximately 9 years comes from table 3-9 of the American Housing Survey National Summary Report and Tables for 1997-2005.

The average floor space and square footage occupied by a renter relative to an owner comes from tables 2-3, 3-3, and 4-3 of the American Housing Survey National Summary Report and Tables for 1997-2007. We compare the median square footage per person and lot size of owners and renters who moved in the last year. The average is 0.637 for square footage and 0.594 for lot size, so we use $\zeta = 0.6$.

The annual persistence of local income shocks of 0.885 comes from Table 1 of Glaeser et al. (2014). They estimate an ARMA(1,1) model for income using Home Mortgage Disclosure Act data from 1990 to 2004. They find AR(1) coefficients of 0.88 for he coastal regions, 0.89 for the sun belt, and 0.88 for interior regions. We average these to get 0.885.

To determine $\gamma_I$, the incidence of income shocks for houses in negative equity, we divide the seasonally adjusted number of foreclosures by the maximum seasonally
adjusted number of homes in negative equity in the CoreLogic data. The mean annual incidence is $\gamma_I = 8.6\%$.

$\sigma$ is calibrated based on bank, FHFA, and Time out of market for foreclosure as described in the main text. The policies for how long a buyer must wait after a foreclosure to be eligible for a mortgage can be found at http://www.nolo.com/legal-encyclopedia/when-can-i-get-mortgage-after-foreclosure.html and http://www.zillow.com/blog/boomerang-buyers-buying-again-after-a-foreclosure-or-short-sale-102457/.

**Loan Balance Distribution Calibration**

We use a minimum-distance methodology to calibrate the loan balance distribution for each geography. From the 7 quantiles given to us by CoreLogic and the Census data on the number of owner-occupied homes without a mortgage, we construct a CDF of 6 points: the fraction of loans with under 50% LTV, under 60% LTV, under 70% LTV, under 80% LTV, and under 100% LTV. We then construct a weighted sum of squared differences between the Beta distribution and the empirical CDF. Because the upper tail of the distribution is most critical for our amplification channel, we weight the under 50%, under 60%, and under 70% parts of the distribution by .1 and the under 80%, under 90%, and under 100% by .2. We then choose $b_a$ and $b_b$, the parameters of the Beta distribution, to minimize this norm. The resulting fit is close enough that our results are robust to alternate weightings.