FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

Housing Supply and Foreclosures

William Hedberg
Federal Reserve Bank of San Francisco

John Krainer
Federal Reserve Bank of San Francisco

September 2012

Working Paper 2012-20
http://www.frbsf.org/publications/economics/papers/2012/wp12-20bk.pdf

Housing Supply and Foreclosures*

September 2012

William Hedberg and John Krainer

Abstract

We explore the role of foreclosure inventories in a model of housing supply. The foreclosure variable is necessary to account for the steep and sustained drop in new construction activity following the U.S. housing market bust beginning in 2006. There is modest evidence that local banking conditions play a role in determining housing starts. Even with state-level foreclosures and banking variables in the model, there is a sizeable post-2006 residual common to all states. We argue that, in addition to observable macro and local factors, housing starts in the Great Recession have been weighed down in part by aggregate uncertainty factors.

JEL Codes: R10, R31, G21, E22

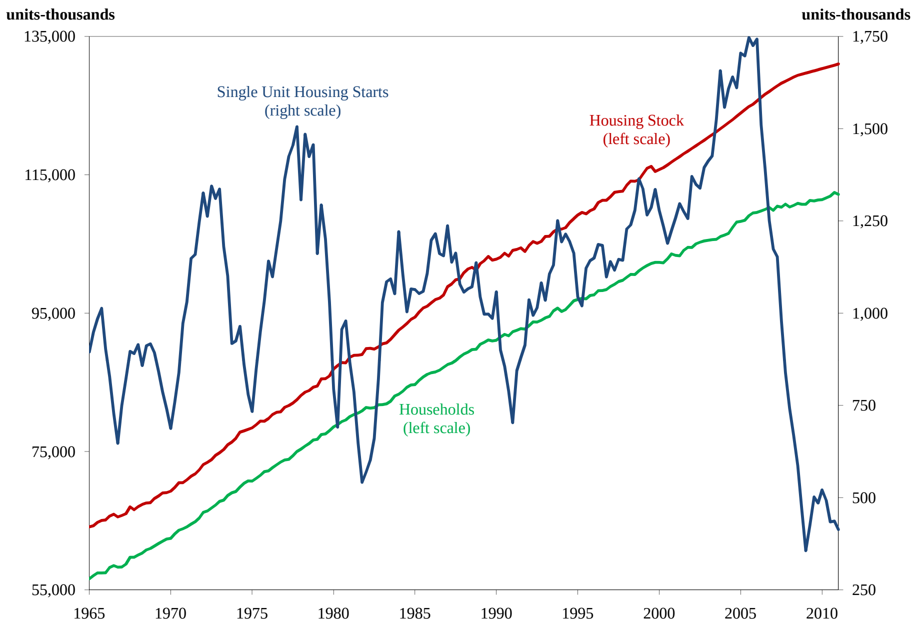
*Both authors are from the Federal Reserve Bank of San Francisco. We thank Fred Furlong and Harvey Rosenblum for helpful comments. The views expressed are those of the authors and not necessarily those of the Federal Reserve System.

1 Introduction

In this paper we update a standard model of the housing supply function to assess how well the model can account for dynamics in the housing sector during the Great Recession beginning in 2007. As figure 1 suggests, the last two decades of housing construction data have been volatile and quite eventful. The residential construction sector enjoyed a sustained boom from the end of the 1990 recession to 2006, after which activity collapsed and then stayed at extremely low levels throughout the ensuing recession. We show that many of the traditional factors that enter into a housing supply function, such as price signals from the product market and changes in costs, simply cannot account for the large and sustained drop in construction activity. When we augment the model to include a variable measuring the foreclosure inventory, we find dramatic improvements in model performance. When we shift the analysis to the state-level data, we provide some evidence that foreclosures appear to represent more than just extra inventory that needs to be worked off before new construction can resume. It is generally true that construction fell more in states where foreclosure rates were higher, but construction activity dropped to historically low levels in virtually every region. Given the great amount of cross-sectional variation in variables like foreclosure rates and house price changes, we conclude from the state-level analysis that foreclosure inventories likely proxy for some deeper underlying uncertainty that has weighed down construction activity at an aggregate level.

The stock of housing changes slowly over the time in order to accommodate population growth and new household formation. While these demographic factors change very slowly over time, over the short run, new additions to the housing stock can be quite volatile (see Figure 1). This is understandable, given the irreversible nature of transactions in the housing market. On the demand side, much of the newly constructed housing in the United States is purchased by first-time homeowners. This class of home buyers can be quite sensitive to changes in the economic outlook for house prices, jobs, financial wealth, and other variables that impact their budget constraints. There is also considerable uncertainty on the supply side. Developers must make forecasts of development costs and the likely selling price of their projects once they have been completed. All of these variables depend on the realizations of future demand conditions as well as the collective actions of all developers who can bring newly developed real estate to market. If uncertainty over these prices and costs increases, there may be a powerful incentive to delay the projects.

Figure 1: Housing Market Supply and Demand Factors



Real estate development has traditionally been viewed as a variant of the classic firm investment problem where new supply depends on expected revenues and costs. Many researchers have built on this fundamental approach by integrating in additional structure provided by information about the physical characteristics of the land as well as theory on how and why urban centers grow. See Topel and Rosen (1988), Mayer and Somerville (2000), and Green, Malpezzi, and Mayo (2005) for examples. In these papers some of the main conclusions (and controversies) center around the relationship between the timing of new construction and signals from the product market. Topel and Rosen (1988) stress the way new construction can show sluggish adjustment because of the interaction of developer's desire to smooth production with convexities in their cost functions. Estimates of price elasticities can also vary widely. Indeed, DiPasquale and Wheaton (1994) show how predictions about how seemingly easy-to-predict variables like demographics rest crucially on what the assumption is about the price elasticity of supply.

Real estate development can also be viewed as an option. See Titman (1985) and Williams

(1991) for theoretical adaptations of the Black-Scholes-Merton option-pricing framework to real estate. See Quigg (1993) and Bulan, Mayer, and Somerville (2006) for empirical implementation. The central insight from this literature is the way that uncertainty interacts with the investment timing decision and the way that this timing, in turn, depends on the competitive environment. This uncertainty feature seems particularly important in the recent episode given weakness of household balance sheets following the financial crisis and the effect of financial market volatility on a finance-dependent sector such as construction.

In this paper we borrower from both strands of the housing supply literature and also incorporate several variables that played prominent roles in the current housing cycle, but less so in previous cycles. Chief among these new variables is the mortgage foreclosure rate. We show that it is very difficult to match the path of housing starts over the last decade without a variable such as this in the empirical specification. Indeed, given the recent behavior of variables generally thought to belong in a housing supply equation, new home construction should have recovered to long-run average levels in 2010.

The fact that foreclosures appear to be a necessary addition to the set of explanatory variables in the housing supply function makes it interesting to study disaggregated data. We do this for two reasons. First, foreclosure laws differ by state, with some states pursuing a judicial foreclosure process whereby lenders must sue borrowers in court in order to repossess the collateral. Other states, by contrast, use a non-judicial foreclosure process where the time-consuming process of getting a judge to approve a foreclosure is avoided. These legal differences have been shown to imply differences in mortgage lending, foreclosure activity, and even house price dynamics (see Mian, Sufi, and Trebbi (2011) and Pence (2006)). Secondly, the foreclosure inventory could proxy for other factors besides the stock of unsold homes on the market. In aggregate, the foreclosure inventory is likely to be correlated with the broader credit market conditions affecting developers and buyers alike. In a pair of recent papers, Duca, Muellbauer, and Murphy (2011a) as well as Duca, Muellbauer, and Murphy (2011b) find an important role for changes in lending standards when accounting for U.S. house price appreciation over the last cycle. In our analysis using disaggregated markets, we can use data on the banks operating in local markets to hopefully disentangle these effects.

The paper is organized as follows. In section 2 we outline the data and empirical methodologies used in the paper. In brief, we adapt the basic setup of Mayer and Somerville (2000) to include measures of the foreclosure inventory as well as expectations for future volatility. In section 3 we apply the models to aggregate data to better understand the role of foreclosures in suppressing new building. Section 4 extends the analysis to the state-level data. Section 5 concludes.

2 Basic theory

As described in DiPasquale and Wheaton (1994), most treatments of the housing supply function are built up from two fundamental equations:

$$D(X, P^*) = S(Z, P^*), \tag{2.1}$$

$$\Delta S = C - \delta S. \tag{2.2}$$

Equation 2.1 specifies housing demand and supply functions that depend on house prices as well as demand and supply shifters X and Z, respectively. There exists a price, P^*, that equates demand to supply in housing market equilibrium. Equation 2.2 is an identity: the change in the housing stock is equal to new construction less the depreciated stock. Most empirical analysis focuses on equation 2.2, taking housing starts as the variable measuring the change in supply, ΔS. Many of the more recent contributions in housing economics have sought to loosen the period-by-period market clearing condition in equation 2.1 and insert sluggish adjustment of the housing stock to changes in the economic environment. These frictions could stem from rising marginal production costs (Topel and Rosen (1988)), regulatory constraints (Malpezzi (1996), Quigley and Raphael (2005)) or physical constraints on new building (Saiz (2010)).

In this paper we seek to incorporate frictions to the housing adjustment process that originate from two relatively underexplored sources. First, we introduce a measure of the foreclosure inventory into the set of covariates (Z in equation 2.1) to provide a sharper characterization of the vacancy rate. Foreclosures could simply reflect vacant or soon-to-be vacant units that must eventually be sold. Under this interpretation, foreclosures act much like other sources of housing inventory in the way they create a wedge between the actual housing supply and the desired supply. Alternatively, the foreclosure inventory could reflect vacancy and also uncertainty about current and future demand conditions. If the number of foreclosures and distressed borrowers is high, developers may take this as a signal that future consumer income is uncertain and sales (and future sales prices) may then be depressed. This type of uncertainty would closely track the type of options-related motive for delaying investment (e.g., Titman (1985) and Bulan, Mayer, and Somerville (2006)).

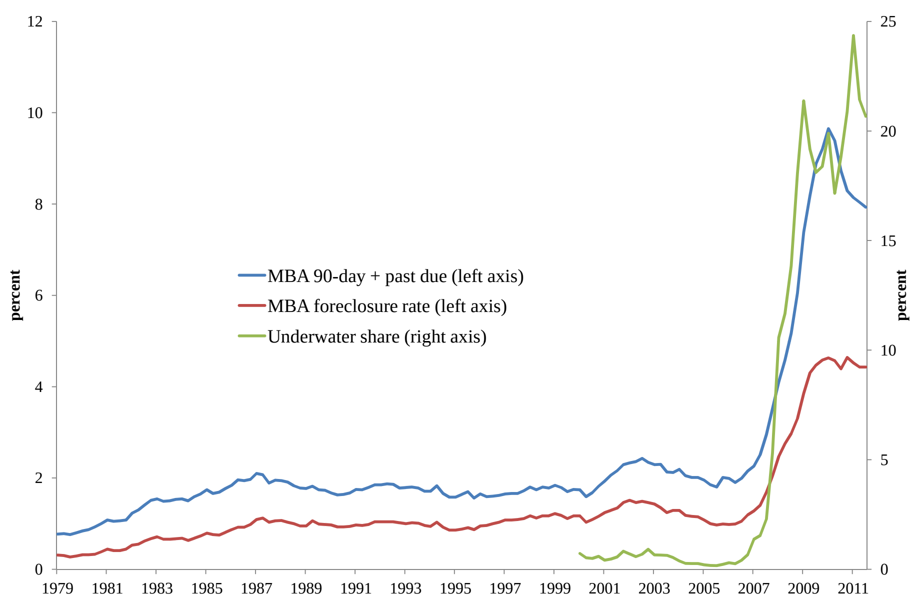
Another channel that we examine is the bank lending channel. In bad times banks perceive their own cost of capital to go up, which should shift the loan supply function and curtail lending. This shift in the supply of credit may not be clearly detectable by simply looking at loan prices. If banks face a loan demand function that is characterized by a differentiation in the riskiness of underlying projects that require finance, then tighter loan terms result in relatively more lending to less risky projects and a drop in risky lending. If this happens, then we would not necessarily observe significantly higher interest rates, but may rather see a change in the type of loans being originated. To get at this notion of shifting credit standards we use a measure of the regulatory concern about banking sector health to proxy for possible changes in bank credit standards. Specifically, we look at the share of banks in the state with a CAMELS rating considered to be unsatisfactory.[1] For each state we construct this measure by aggregating all the CAMELS ratings of banks headquartered in that state. This aggregation method obviously leads to some inaccuracies: for example, a large bank such as Wells Fargo's rating would contribute to California's measure of bank health and no other, even though Wells Fargo lends in a large number of states. However, since construction loans are thought to be one of the few remaining relationship-intensive form of lending, our requirement that a bank needs to have its headquarters in the state in order to lend there is probably helping us to detect institutions that are the marginal lenders in real estate development.

[1]The CAMELS rating is a summary measure of bank condition that is assigned following an on-site bank examination. The examination team evaluates commercial bank capital plan ("C"), asset quality ("A"), managerial processes ("M"), earnings robustness ("E"), liquidity position ("L"), and sensitivity to market conditions ("S"). Ratings are on a scale of 1-5, with ratings in the range of 3-5 indicating that there are matters of supervisory concern pertaining to that institution.

3 Data and Empirical Methodology

The housing supply function, measured here by single-family housing starts, is modeled as a function of revenue and cost variables. Following Mayer and Somerville (2000) we adopt a specification where starts depend on real house price changes and not the level of real house prices, as in Topel and Rosen (1988).[2] This specification ensures that temporary spikes in house prices, like would be seen if a market experienced a one-off demand shock, result in temporary and not permanent increases in construction activity. Topel and Rosen (1988) emphasize how convex adjustment costs of developers and a desire to smooth production can potentially result in delayed responses of starts to house price changes. Thus, in all of the specifications that we try here we include at least three lags of the house price change. We use the Freddie Mac House Price Index, primarily for the virtue that its history extends back to the 1970's and disaggregated versions of this series are also available.

Figure 2: Measures of the Inventory of Distressed Homes



Our foreclosure variable is taken from the Mortgage Banker's Association (MBA), which covers roughly 85 percent of the market. We use the foreclosure rate in our analysis, which is defined as

[2]We conducted Dickey-Fuller tests on both starts and house price changes, and found no evidence of nonstationarity.

the total inventory of homes in foreclosure as a percentage of total mortgages.[3] As can be seen in figure 2 the recent dynamics of the MBA foreclosure index are quite similar to an alternative index of mortgage distress such as the share of "underwater" mortgages with principal balance greater than estimated house price.[4] We use the MBA index because of the longer period of availability. The interquartile range of this measure of the foreclosure rate is fairly tight, between .88 and 1.19 percent. However, the most notable feature in 2 is the dramatic upward swing in the foreclosure rate to nearly 5 percent at the end of the sample period. This upturn, it will turn out, corresponds with the sharp decline in housing starts and will make this variable useful in the starts equation that we estimate.

Table 1: Summary Statistics

	mean	stdev	25th pctile	median	75th pctile
1-unit housing starts (000s)	276.1	74.7	246.0	282.2	319.2
Real house price change	130.8	1128.6	-271.3	163.3	770.0
Housing stock (000s)	105,891.6	15,498.5	92,115.8	105,812.9	118,733.9
Months for sale	5.56	2.01	4.20	5.03	5.89
Foreclosure stock (000s)	369.0	486.8	96,166.4	180,090.8	435,075.3
Foreclosure rate (percent)	1.24	.94	.88	1.02	1.19
Real change materials cost	-.027	.491	-.317	-.102	.214
Real change prime rate	.000	.906	-.296	.012	.303
Real change bank rate	.002	.986	-.309	.015	.261
Low CAMELS share	.114	.077	.051	.072	.196
Change in low CAMELS share	8.42	21.11	-5.3	1.9	16.1
Number of observations	144	144	144	144	144

Our main specification takes the following form,

$$s_t = \alpha_0 + \gamma H_{t-1} + \beta F_{t-1} + \sum_j \theta_j \Delta p_{t-j} + \rho \Delta c_{t-1} + \sum_k \phi_k \Delta b_{tk} + \epsilon_t. \tag{3.1}$$

H_{t-1} is the lagged stock of housing, so the coefficient γ represents the steady-state replacement rate of structures. F_{t-1} denotes the level of the foreclosure inventory. As noted above, house prices (and their lags) enter into the estimating equation in differenced form, Δp_{t-j} for $j = 1, ..., J$. We also consider one lag of construction costs, also differenced, Δc_{t-1}, and then a vector of credit-related variables, b_{tk}.

[3] That is, the foreclosure rate is relative to the stock of outstanding mortgages and not relative to the total number of households.

[4] The underwater mortgage share is based on data from LPS Applied Analytics

Most studies of the housing supply function employ some sort of instrumental variables strategy to control for the endogeneity of starts and house prices. The endogeneity problem is a standard omitted variables problem; there could be a variable that is driving both starts and house prices that is not controlled for in the regression. Further, there is a possibility of reverse causality from starts to house prices: presumably, a large influx of new construction onto the market will have a depressing effect on house prices. To control for these possibilities we estimate a first stage regression of house price changes on lagged demand factors, including changes in mortgage interest rates, the number of married couples, changes in nonfarm employment, the user cost, as well as lagged values of all the exogenous variables in the specification above. Similarly, we also instrument for construction costs using variables such as changes in energy prices, its own lags, and the lags of the other exogenous variables. As was the case for house price changes, here we are primarily concerned with reverse causality where construction activity (or lack thereof) affects construction input prices.

The results from the first-stage regressions can be found in Tables 2 and 3. Most of the variables in the first-stage price change regressions have the expected signs. Employment growth is associated with house price appreciation. Positive changes in mortgage rates and energy prices, which tend to occur during good economic times, are positively associated with increases in house prices. There are many interesting differences that emerge when this specification is estimated on a short sample (1987.Q3 - 2005.Q4) as compared to the full sample (1987.Q3 - 2010.Q4). In particular, the coefficients on months for sale and the real user cost switch sign across sample periods. The change in CAMELS ratio comes in significant only in the longer sample period, as we might expect given the large number of downgrades that took place during the 2007 recession. Note also that lagged foreclosures are positively associated with house price changes in the 1987-2005 sample, suggesting a channel running from house price appreciation to decreased affordability, and hence to foreclosures.

In general, the instruments used in Table 2 are quite weak. Only the user cost comes in significant in the 1987-2010 sample. The demographic variables–change in married couples–have the right sign, but are not estimated with precision. The instrument for changes in construction

Table 2: First stage regression for house price changes

	1987.Q3 - 2005.Q4	1987.Q3 - 2010.Q4
Change in employment	.687	.970
	(.244)***	(.459)**
Change in employment(-1)	.172	-.932
	(.285)	(.436)**
Change in energy prices	4.177	21.674
	(11.204)	(12.856)*
Change in energy prices(-1)	2.886	22.193
	(10.914)	(13.276)*
Change real mortgage rate	138.994	484.282
	(131.855)	(303.858)
Change real mortgage rate(-1)	-228.523	165.011
	(127.916)*	(292.026)
Change in married couples	-.988	1.168
	(.729)	(1.367)
Change in married couples(-1)	1.803	.803
	(.717)	1.323
Housing stock (-1)	.027	-.139
	(.029)	(.039)***
Months for Sale(-1)	199.513	-454.384
	(93.726)**	(131.003)***
User cost	340.051	-924.056
	(110.520)***	(197.105)***
Foreclosure(-1)	8.630	-.195
	(1.571)***	(.930)
Change CAMELS ratio	6433.217	75472.58
	(9461.431)	(21257.36)***
Constant	-9166.457	26326.4
	(3999.108)**	(6080.366)***
R^2-adjusted	.832	.538
Number of observations	74	94

materials costs is the lagged change. In Table 3 we see that the lag is significantly correlated with changes in materials costs. Note that none of the other variables that will enter into the second stage (housing starts) regression come in significant in Table 3.

Table 3: First stage regression for construction materials cost changes

	1987.Q3 - 2005.Q4	1987.Q3 - 2010.Q4
Change in construction materials cost(-1)	.441	.494
	(.128)***	(.096)***
Housing stock(-1)	-2.41E-6	-3.54E-6
	(2.08E-5)	(1.31E-5)
Months for sale	-.029	-.066
	(.070)	(.061)
Foreclosures(-1)	.0007	.0004
	(.001)	(.0004)
Change real prime rate(-1)	.098	-.128
	(.127)	(.108)
Change CAMELS ratio(-1)	-3.017	-8.567
	(8.032)	(8.483)
Constant	.289	.604
	(2.115)	(1.617)
R^2-adjusted	.225	.266
Number of observations	74	94

The main results for the aggregate supply function may be found in Table 4. In all of the specifications looked at here, the most recently observed change in house prices is positive and economically significant. The effects of house price changes wane fairly quickly, as the longer lags are rarely significant. The supply elasticity with respect to house prices is in the range of .02 to .06. These estimates are on the low end of the range produced in the empirical literature, but in line with the results in Mayer and Somerville (2000).[5]

Few of the other cost shifters in Table 4 come in significant. The signs on many of these variables (e.g., the change in the interest rate, the change in real construction materials costs) are positive–the opposite of what one would expect. The share of banks with low CAMELS ratings has a negative effect on starts, although the effect is not statistically significant. This somewhat

[5]Other studies cited in the references produce higher estimates of the price elasticity, in the neighborhood of 1.0 to 2.0, probably due to the different sample periods used for the analysis and different functional form assumptions.

puzzling result is also consistent with the empirical literature on housing supply. Either costs have not varied enough over time to trigger changes in building, or developers are largely able to pass on cost increases to eventual buyers. In either case, in this specification that relies on time-series data to identify the housing supply function, price signals seem to be paramount in influencing building decisions.

Despite the insignificance of many variables that theory would suggest as being relevant for building decisions, the model appears to be able to capture the time-series properties of the starts data fairly well. One of the most interesting results here is how important the foreclosure inventory is for fitting the aggregate data. Many of the coefficient estimates change quite dramatically from specification (i) (no foreclosures) to specification (ii) (foreclosures included). The model R^2 leaps from about 51 percent in specification (i) to about 84 percent in specification (ii). This point can be made most forcefully by plotting the in-sample model fit for the two specifications. In Figure 3 we see that the "standard model" without foreclosures is simply incapable of fitting the housing starts data over the last several years of the sample. Indeed, with the cut in interest rates and the stabilization of house prices after the recession was in full-swing, this model actually signals a strong surge in housing starts that never came about.

4 State-level estimates of the supply function

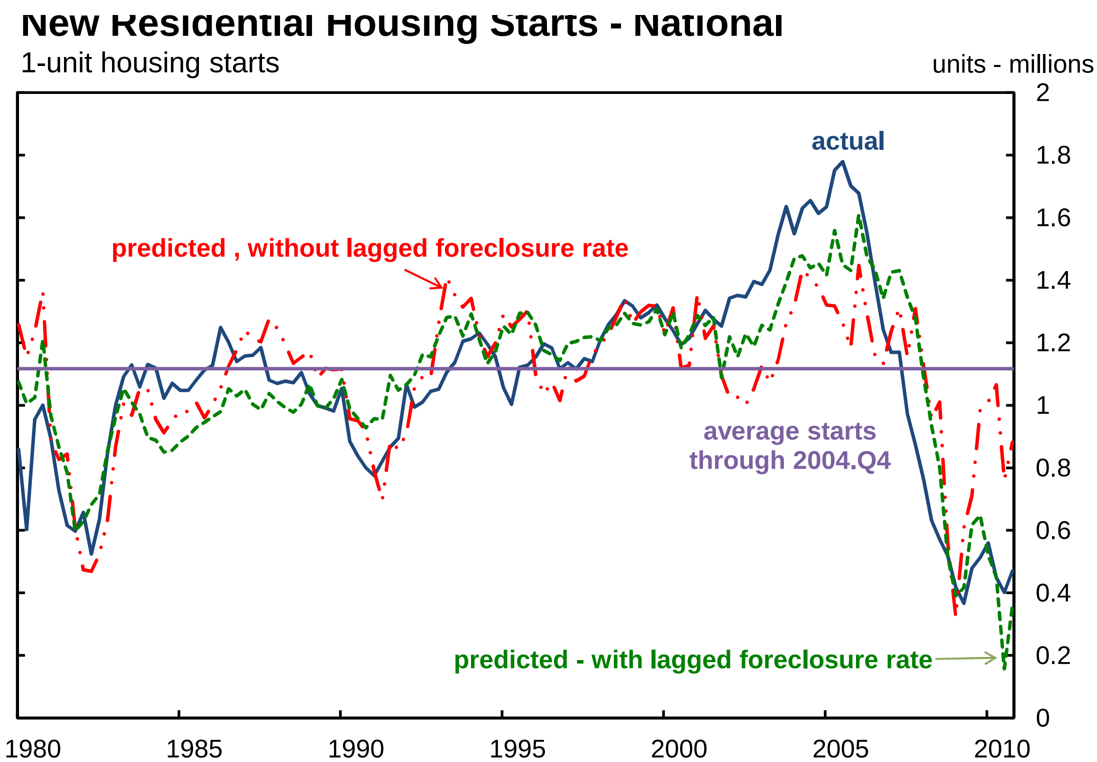
One distinct difficulty in estimating the housing supply function is the relative paucity of time-series data for the task. Figure 2 is a reminder that, insofar as studying the housing cycle, researchers essentially have just one episode from which to observe how new home construction is related to mortgage default and homeowner distress. This shortcoming leads us to consider the cross-sectional variation in housing supply over the cycle.

There is now a large body of evidence suggesting that housing market functioning and outcomes differ substantially across geographic regions. Green, Malpezzi, and Mayo (2005) estimate supply elasticities at the Metropolitan Statistical Area (MSA) level and report values ranging from negative values up to nearly 20 for the most price-elastic markets. Saiz (2010) also finds considerable cross-sectional variation in key elasticities that he attributes to physical constraints on new building (e.g.,

Table 4: Aggregate Model of Housing Starts

	(i)	(ii)	(iii)	(iv)	(v)
Change in real house price	.056	.031	.030	.031	.027
	(.013)***	(.013)**	(.007)***	(.007)***	(.006)***
Change in real house price(-1)	.006	.008	.010	.010	.009
	(.016)	(.009)	(.007)	(.007)	(.006)
Change in real house price(-2)	-.001	.012	.012	.012	.011
	(.017)	(.010)	(.008)	(.008)	(.007)
Change in real house price(-3)	.031	.013	.013	.014	.014
	(.013)*	(.009)	(.007)	(.007)	(.006)*
Real change materials cost	71.596	1.437	20.192	21.419	23.073
	(27.841)*	(.432)*	(15.158)	(15.335)	(14.809)
Change real prime rate	.443	.707	8.530	8.721	7.569
	(17.216)	(.355)	(9.003)	(8.939)	(8.168)
Change real prime rate(-1)	-.633	.914	-3.622	-4.026	-5.119
	(16.941)	(.357)*	(8.794)	(8.782)	(8.153)
Foreclosures(-1)		-1.818	-46.268	-47.626	-37.285
		(9.504)***	(.171)***	(12.656)***	(12.620)**
Housing Stock(-1)				.003	.003
				(.001)***	(.001)***
Time trend	.260	1.205	1.200		
	(.256)	(.248)***	(.234)***		
Months for Sale(-1)			3.980	4.806	2.446
			(4.805)	(4.888)	(4.396)
Low CAMEL share					-944.30
					(580.481)
Constant	244.741	219.073	251.116	-248.368	-182.129
	(26.093)***	(16.609)***	(21.191)***	(121.925)*	(112.308)
R^2-adjusted	.507	.810	.863	.863	.886
Log Likelihood	-492.852	-448.888	-433.528	-433.438	-424.631
Number of observations	91	91	91	91	91

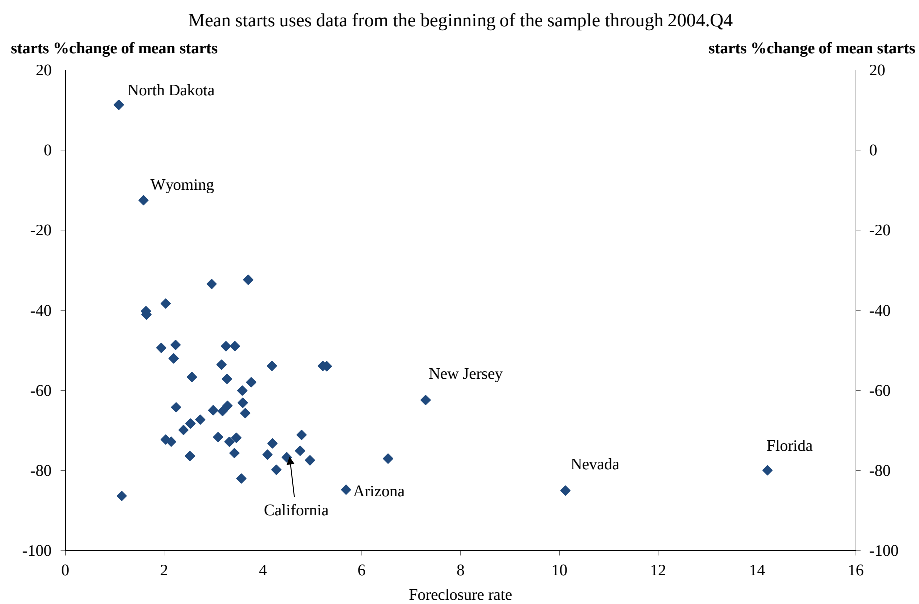
Figure 3: Fitted values of the housing supply function



mountains, bodies of water).

Especially important for our analysis is the geographic variation that arises from the uneven pattern of foreclosure rates across the country. In figure 4 we show that, while the negative relationship between starts and foreclosures is still apparent in the data, construction activity during the 2007 recession was much lower in virtually all states relative to their pre-2004 average.[6] Thus, further analysis needs to be done to differentiate between a story where foreclosures weigh on starts through a pure inventory effect, as opposed to high foreclosure rates acting as a proxy for general problems in the credit markets or spillovers into overall housing demand.

Figure 4: State-level foreclosure rates and housing starts



As an exploratory analysis we can extract the main principal components from the individual state-level housing starts series to get a gauge for how much common variation there is in new home construction. The specifications analyzed the aggregate starts data (Table 4) would suggest a fairly strong role for local economic variables, given the importance of variables such as house prices changes and foreclosure rates. Indeed, prior to the housing peak in 2006 we see the first

[6]We conducted Dickey-Fuller tests on the state-level starts series and found all of them to be stationary. Thus, figure 4 plots deviations of starts from the states' long-run averages and not from the trend.

principal component of the explains roughly 71 percent of the common variation in the housing starts activity at the state level (column 2 of Table 5. Interestingly, the degree of common variation in starts leaps once housing market stalls and the recession begins. Housing starts fell almost everywhere to all-time low levels. This observation will turn out to be important when interpreting the supply function estimates from the disaggregated data. In spite of fairly different housing market conditions at the regional level, actual housing starts moved together in lock-step for a large part of the recession period.

Table 5: Principal components analysis of state-level housing starts

	full sample	pre-2006	post-2006
First principal component	0.88	0.71	0.94
Second principal component	0.92	0.84	0.98
Third principal component	0.95	0.90	0.99
Observations	129	110	19
States	50	50	50

In all of our specifications using the state-level data we include a rolling 8-quarter standard deviation of state-level house price changes, as well as an indicator variable for whether the state law calls for a judicial or nonjudicial foreclosure.[7] We also include state fixed effects. In Table 6 we report the results from the state-level regressions. The data are pooled. As was the case with the aggregate starts data, we instrument for house price changes and changes in construction materials using demographic variables for house price changes and lagged values. We estimate the supply function in levels. Thus, the estimated coefficients for many of the covariates are not directly comparable given the differences in scale at the state level. Still, many of the same features from the aggregate data come through here as well. The most recent lag of real house price changes is strongly significant, demonstrating again that developers pay close attention to price signals before making their building decisions.

Lagged foreclosures still has an important role explaining fluctuations in starts, as was the case in the aggregate data. Interestingly, we do not see much evidence in Table 6 of bank effects on

[7] In a judicial state the lender must conduct a foreclosure through the court system, which is generally thought to be slower and more costly, giving rise to longer foreclosure periods all other things held constant.

housing activity. Only the level of real estate chargeoffs in the state has a significant negative impact on starts. The state-wide share of low CAMELS banks fails to come in significant in any of the specifications.

Many of the stylized results from the analysis of aggregate starts carry over to the state-level analysis as well. Namely, without the foreclosure variable in the regression, it is not possible to match the precipitous decline in starts post-2006, nor can we justify the persistent low level of starts that endures to the present day. However, we can also observe that the improvement in overall model fit gained from adding the foreclosure variable is not nearly so dramatic as was the case for aggregate starts. In columns (ii) and (iii) of Table 6 the regression R^2 goes from 80 percent to about 83 percent when adding the local foreclosure rates. Thus, it seems apparent that some form of uncertainty is weighing on the housing market that is better captured by foreclosures at the aggregate level than at the state level. This suggests that foreclosures affect starts not just because a house in foreclosure represents an addition to the inventory of units of sale, but because of the uncertainty or heightened risk aversion that seem to track foreclosures at the aggregate level.

In column (vi) of Table 6 we include a specification where we explicitly control for spatial effects that might be present in the data. For example, if the shocks to the housing market are regional in nature, then particularly hard-hit regions will tend to have similar outcomes. Failure to control for this could potentially yield inconsistent parameter estimates. The specification (vi) of Table 6 is from the class of "spatial AR" models where we allow the error term in the starts equation to be serially correlated and dependent on the lagged errors of states that share a boundary. With some simplifications to the notation used previously, the vector s_t collects starts in all states, and is assumed to be a function of the matrix X, which contains the explanatory variables that we have discussed already, some if which vary at the state level. The spatial AR model is

$$s_t = \rho W s_t + \beta X_t + u_t, \tag{4.1}$$

where W is a weighting matrix with element $w_{ij} = 1$ if states i and j are neighbors, and $w_{ij} = 0$ otherwise. The scalar parameter ρ and is the spatial correlation parameter that measures the dependence of starts amongst neighboring states.

There does appear to be significant spatial dependence in the starts data. With the point estimate for $\rho = 0.41$ and strongly significant, there is some omitted variable related to starts that varies quite strongly amongst neighbors. Before discussing the implications of this more fully, however, note that there are several other differences between the point estimates from column (vi) compared to (i)-(v) of Table 6. The coefficient on lagged foreclosures is much larger in magnitude after having controlled for the spatial correlation. Additionally, the coefficient on judicial foreclosure variable is now positive and significant. All else held constant, less lender-friendly foreclosure laws lead to a lower level of housing starts.

We return to the presence of spatial dependence in the housing starts data, which we find to be significant and positive. Conditioning on observables like state-level house price changes, lagged foreclosures, and the other variables, there is still a correlation between starts in neighboring states that is not being captured by the model. This type of finding is often interpreted in the literature as evidence of regional economic integration that does not necessarily stop at a state's borders. In a model with state-level variables and aggregate variables, this case would give rise to spatial dependence like what we find here. But in our case, there appears to be an interesting form of time variation in the role played by spatial effects. We can see this in Figure 5 where we plot how the cross-section of state residuals from the starts equation varies over time. Once the housing bust set in around 2006, the model tends to over-predict housing starts on average (i.e., the residual is negative). The standard deviation of the cross-section of residuals actually falls over this time as well (the dashed line in Figure 5). Thus, even controlling for foreclosures and for regional dependencies, we still have difficulty accounting for the steep decline of starts during the housing market crash.

5 Conclusion

We find an important role for the mortgage foreclosure rate in explaining the dynamics of housing starts following the housing boom ending in 2006. There is some evidence that constraints on banks also played a role in restraining new supply. However, in terms of economic significance, a variable like foreclosures is vital for accounting for both the drop and the sustained depression in

Table 6: State-Level Model of Housing Starts

	(i)	(ii)	(iii)	(iv)	(v)	(vi)
Change in real house price	.002 (.0003)**	.001 (.0003)**	.002 (.0002)***	.001 (.0003)**	.001 (.0002)***	.001 (.0003)***
Change in real house price(-1)	.0005 (.0002)*	.0009 (.0002)**	.001 (.0002)**	.001 (.0003)*	.0008 (.0002)**	.0005 (.0003)
Change in real house price(-2)	.0005 (.0003)	.0006 (.0002)*	.001 (.0002)**	.001 (.0004)	.0006 (.0002)**	.0006 (.0003)*
Change in real house price(-3)	.0004 (.0004)	.0002 (.0002)	.001 (.0002)**	-.0004 (.0003)	.0003 (.0002)	.0005 (.0003)
State HPI st.dev.				-.0001 (.0002)	-.0001 (.0002)	-.0001 (2.48E-5)***
Real change materials cost	2.140 (.762)*	1.259 (.508)*	.844 (.842)	.890 (.988)	1.842 (0715)*	.475 (.608)
Change real prime rate	.374 (.232)	.909 (.343)*	.568 (.378)	.850 (.326)*	.822 (.333)*	.412 (.323)
Change real prime rate(-1)	.374 (.232)	-.138 (.156)	-.045 (.153)	-.136 (.147)	-.056 (.223)	.351 (.318)
Foreclosures(-1)	-.072 (.015)**	-1.836 (.316)***		-.089 (.013)***	-1.899 (.291)***	-.45 (.109)***
Judicial Foreclosure	-4.833 (.113)***	-5.301 (.22)***	-4.723 (.582)***	3.205 (3.499)	-1.19 (4.418)	-.499 (.156)***
Housing Stock(-1)				5.75E-6 (2.4E-6)*	2.7E-6 (2.89E-6)	3.11E-6 (3.47E-8)***
Time trend	.041 (.025)	.043 (.025)	-.004 (.014)			
Months for Sale(-1)		-.032 (.154)*	-.191 (.097)	-.464 (.149)	-.037 (.267)	.24 (.034)***
Real Estate Chargeoffs		-1.7E-6 (7.44E-7)*				
Equity capital			1.4E-7 (2.27E-7)			
Low CAMEL share		-27.513 (39.311)	-96.779 (37.786)*		-56.699 (42.049)	-18.41 (9.92)
Lagged spatial error						0.415 (.016)***
R^2-adjusted	.831	.837	.809	.856	.837	
Log Likelihood	-14,099.02	-14,009.6	-14,373.71	-13,738.3	-14,007	
State fixed effects	yes	yes	yes	yes	yes	no
Number of observations	4,500	4,500	4,500	4,500	4,500	4,500

Figure 5: Residual Diagnostics from Spatial Model of Housing Starts

new home construction. While this main result holds up when we expand the analysis to include data from the states, there remains evidence that aggregate factors have influenced starts, beyond those factors traditionally used in a housing supply equation. The increases in foreclosure rates during the recession were more uneven across states. While foreclosure rates generally improve our ability to fit the state-level starts data, the improvement is not nearly so dramatic as is the case in the aggregate. Yet housing starts fell virtually everywhere by more than our model would predict. This suggests that there is some form of uncertainty still hanging over the housing market that is correlated with the aggregate foreclosure rate, but is weighing on the market in a different way than local inventory that needs to be cleared.

6 References

Bulan, L., C. Mayer, and C. T. Somerville (2006): "Irreversible Investment, Real Options, and Competition: Evidence from Real Estate Development," *Journal of Urban Economics.*

DiPasquale, D., and W. C. Wheaton (1994): "Housing Market Dynamics and the Future of Housing Prices," *Journal of Urban Economics*, 35.

Duca, J. V., J. Muellbauer, and A. Murphy (2011a): "House Prices and Credit Constraints: Making Sense of the U.S. Experience," Federal Reserve Bank of Dallas working paper.

——— (2011b): "Shifting Credit Standards and the Boom and Bust in U.S. House Prices," Federal Reserve Bank of Dallas working paper.

Green, R. K., S. Malpezzi, and S. K. Mayo (2005): "Metropolitan-Specific Estimates of the Price Elasticity of Supply of Housing, and Their Sources," *American Economic Review Papers and Proceedings*, 95(2), 334–339.

Malpezzi, S. (1996): "Housing Prices, Externalities, and Regulation in U.S. Metropolitan Areas," *Journal of Housing Research*, 7(2), 209–241.

Mayer, C. J., and C. T. Somerville (2000): "Residential Construction: Using the Urban Growth Model to Estimate Housing Supply," *Journal of Urban Economics*, 48, 85–109.

MIAN, A., A. SUFI, AND F. TREBBI (2011): "Foreclosures, House Prices, and the Real Economy," NBER working paper.

PENCE, K. M. (2006): "Foreclosing on Opportunity: State Laws and Mortgage Credit," *Review of Economics and Statistics*, 88, 177–182.

QUIGG, L. (1993): "Empirical Testing of Real Option-Pricing Models," *Journal of Finance*, 48(2), 621–640.

QUIGLEY, J. M., AND S. RAPHAEL (2005): "Regulation and the High Cost of Housing in California," *American Economic Review Papers and Proceedings*, 95, 323–328.

SAIZ, A. (2010): "The Geographic Determinants of Housing Supply," Forthcoming in the Quarterly Journal of Economics.

TITMAN, S. (1985): "Urban Land Prices under Uncertainty," *American Economic Review*, 75(3), 505–514.

TOPEL, R., AND S. ROSEN (1988): "Housing Investment in the United States," *Journal of Political Economy*, 96(4), 718–740.

WILLIAMS, J. T. (1991): "Real Estate Development as an Option," *Journal of Real Estate Finance and Economics*, 6(4), 191–208.