A Search for the Underlying Structure
Driving House Prices in a Distressed Environment

James R. Follain

© 2012 Lincoln Institute of Land Policy

Lincoln Institute of Land Policy
Working Paper

The findings and conclusions of this Working Paper reflect the views of the author(s) and have not been subject to a detailed review by the staff of the Lincoln Institute of Land Policy.

Contact the Lincoln Institute with questions or requests for permission to reprint this paper. help@lincolninst.edu

Lincoln Institute Product Code: WP12JF2

Abstract

This is the third of three papers prepared for the Lincoln Institute of Land Policy about the housing price bubbles and busts in the US. The first two by Follain and Giertz (2011, 2012) use lengthy time-series data for over 300 MSAs to build models with the potential to predict what actually occurred in the last several years. This paper focuses upon the period 2005–2010, which includes the last years of the boom and the several years of its aftermath. County level data are used instead of MSA level, which offers the opportunity to take a more geographically granular look at the crisis. Annual data for 416 counties are available all six years; data for 439 counties are available for 2006 through 2010. Counties from 47 states and the District of Columbia are included. These data draw upon the American Community Survey, information provided by Collateral Analytics, Moody's Analytics, and other sources and offer the opportunity to study a wide range of potential drivers of house prices during the crisis.

A central goal of this paper is to shed light on the key drivers of house prices during a period of substantial turmoil and distress. Indeed, a number of substantial insights are produced. First among these is the importance of incorporating a measure of the distressed real estate inventory; a larger inventory of such real estate reduces the level of house prices and slows down the growth in house prices. The model also confirms the lingering impacts of excessive price appreciation and substantial amounts of subprime mortgage lending, during the boom and prior to the bust The estimated models also confirm the importance of a number of the traditional drivers of house prices like income and rent, though the empirical magnitudes of these estimates are sensitive to a variety of conditions. The analysis also confirms the presence of large and widely varying county fixed effects, which highlights the importance of local and difficult to quantify market conditions as a driver of house prices. Both the in-sample and out of sample predictions of the model for 2011 suggest that a full recovery is well into the future and also indicate a substantial number of outlier counties that have large and unexplained residuals both positive and negative.

About the Author

James R. Follain is principal of James R. Follain LLC, senior fellow at the Rockefeller Institute of Government at the University of Albany, an advisor to FI Consulting, and a senior faculty member at the Lincoln Institute of Land Policy. He can be contacted at jfollain@nycap.rr.com.

Acknowledgements

Financial support for this project was provided by the Lincoln Institute of Land Policy. Greg Ingram and Yu-Hung Hong of the Lincoln Institute provided substantial input to the design of this research and many helpful suggestions on earlier drafts. We wish to acknowledge the support of Michael Sklarz, the President of Collateral Analytics (http://www.collateralanalytics.com) and his staff, and to Mark Zandi and Celia Chen of Moody's Analytics (http://www.economy.com/default.asp) for providing us with access to some of their wonderful data for this project. Seth Giertz provided valuable feedback and many helpful comments on an earlier draft. Barbara Follain provided valuable editorial assistance throughout the project. Other helpful comments were provided by Brent Ambrose, Wayne Archer and Ling Peng. All errors and omissions are my own.

Table of Contents

A Search for the Underlying Structure Driving House Prices in a Distressed Environment

Section I: Introduction

Economists typically define a price bubble as a level of prices in excess of what seems to be justifiable or in sync with the current and expected levels of the primary drivers of house prices over the long run. For example, a substantial increase in the average price to income ratio over a relatively short time period in a market may be a sign of an emerging bubble. Bubble detection depends on a firm understanding of the drivers of the equilibrium level of prices. Substantial and prolonged positive deviations between the actual level of prices and the levels thought to represent the equilibrium values are suggestive of a bubble and an impending bust. Economists use a variety of structural relationships or criteria to depict the equilibrium level of house prices in the field of bubble detection. This paper seeks to learn more about these underlying structural relationships in the current and severely distressed environment.

One common approach is to measure the equilibrium level of prices by an equation that links the level of house prices to a set of variables thought to be the primary or core drivers of the level of prices. These core variables often include the levels of household income, population, and employment in the market. Examining the relationship between the level of house prices and this set of core variables is one example of what is meant in this paper by the underlying structure of the housing market. Economists often refer to this example as a reduced form equation for the level of house prices. Much of the debate about the potential of a house price bubble involves the selection of the core set of variables to be included in the reduced form equation and its stability over time and among markets.

Besides the reduced form approach, economists consider a variety of other and simpler ways to depict and measure the equilibrium level of house prices in the search for potential house price bubbles. A classic example of this approach focuses upon the rent to value ratio (R/V). In competitive markets with full information, R/V is often expressed as the sum of four components: R/V = interest rate + expenses + risk premium – expected appreciation. If the observed R/V ratio in a market is deemed to be too low relative to what the sum of these four components suggests, then it is often inferred that investors in housing are likely relying upon excessive appreciation in house prices in order to earn a normal rate of return on the investment. Another example of this approach rests upon the price to income ratio (P/Y). If this ratio is deemed to be unjustifiably high, then the potential of a bubble is thought to be high. Other examples of this approach are discussed below.

Still another approach involves the estimation of a system of equations. The system of equations consists of a first stage vector-error correction equation (VEC), which is essentially a reduced form equation for the level of house prices. Other equations in this system focus upon the growth in house prices and other relevant variables such as income per capita and employment. These other equations are referred to as vector autoregression equations (VAR). They incorporate residuals from the first VEC equation as a potential bubble indicator. The VAR system also seeks to capture the degree to which lagged growth in house prices play an important role in the growth of prices. This is referred to as the momentum effect: evidence of larger momentum effects can be another indicator of a potential bubble.

Two previous papers by Follain and Giertz (2011 and 2012) sought to detect house price bubbles and predict future house price growth by using this approach and data at the MSA (Metropolitan Statistical Area) level for long periods of time. Their results showed substantial evidence that the residuals from the first stage VEC affected the growth of house prices as well as substantial momentum effects in the years preceding the bubble bust of the last several years. The predictions of the model varied substantially among markets and with the time periods used to estimate the model. One stylized fact emerged from the first paper, which used data from 1980 through 2010: the ranking of the predictions using data through 2007 were generally in line with the rankings of the actual out of sample price outcomes for these MSAs in years 2008 thru 2010. However, these predictions tended to understate substantially the declines in prices that actually occurred in the worst hit areas such as the SAND states of Arizona, California, Florida, and Nevada.

The current paper shifts focus to the years since the house price bubble began to burst and the distress associated with the bubble bust became most apparent: 2005 through 2010. Its ultimate goal remains the same as in the two earlier papers by Follain and Giertz: estimate models that capture the underlying structural relationship between house prices and their core drivers in order to shed light on where future prices might be headed and to identify those local markets that are far from equilibrium. Presumably house prices are less inflated today than in the years in which the bubble emerged, but perhaps they have overshot the mark and are well below reasonable notions of equilibrium. If so, then above average price growth may be needed to return to equilibrium. Alternatively, perhaps a new equilibrium level of prices has emerged that calls for a substantially lower level of prices than would have been thought appropriate just prior to the emergence of the bubble. A period of prolonged and seemingly depressed level of prices may be on the horizon in this case, especially in those markets in which the bubble and bust were the strongest. The focus of this paper is upon 439 counties for 2005 through 2010. Annual data on well over 100 variables are collected for each county and each year. These counties are located in 47 states and the District of Columbia.

This focus upon county level data during a period of substantial distress offers both unique opportunities and challenges. Using county data provides an opportunity to take a more geographically granular look at housing markets than can be obtained using MSA data. Focusing on this more recent period also permits the utilization of a much richer and wider array of measures of both house prices and the drivers of house prices than can be obtained in a study that focuses upon a longer time-series of relatively stable (healthy) prices, e.g. 1980–2010. These additional data include data from the American Community Survey (ACS), data from a developer of proprietary price indexes (Collateral Analytics), and a wide variety of potential indicators provided by Moody's Analytics. The most important potential opportunity is that focusing upon this period provides an opportunity to study a large number of housing markets in the midst of serious distress and generate potential insights not provided by data dominated by relatively stable market conditions. An analogy to medical research may be helpful to convey this notion. Medical research is likely to identify very different relationships between individual health and diet among people who are generally healthy than among people who are generally unhealthy. If research focuses upon only healthy people the results of the research are likely to be skewed; hence, research also needs to focus upon the unhealthy to understand the relationship between diet and health among the complete population. Similarly, focusing research in this paper on the current period of distress offers an opportunity to broaden the range of market experiences used to understand the drivers of house prices.

The first and generic challenge, of course, is deciphering stable structural relationships and insights about the nature of what the "new normal" may be using data from these deeply distressed markets in which house prices may be far from their equilibrium values. For example, choosing the best set of variables to drive the model may be different than the usual suspects in a study of markets at or near equilibrium. One list of these usual suspects noted above—income, population, and employment—may not be as strongly related to the level of house prices in a distressed environment as they might be in a market peacefully in equilibrium.

A second major challenge facing a search for a stable structural model of the housing market during turbulent times is that the current conditions have been greatly affected by conditions just prior to this bust. Three of these precursors to the bust are highlighted in this study. One is the prevalence of subprime lending in the years immediately before the bust, which varied widely among areas in our data. Home Mortgage Disclosure Act (HMDA) data are used to capture this precondition by focusing on the share of all HMDA mortgage lending deemed to be high cost by HMDA in 2004 and 2005.[1] A second relevant precursor was the run up in house prices prior to the bust, which is proxied by the growth in house prices between 2000 and 2005 using the Federal Housing Finance Agency (FHFA) house price indexes. A third potential precursor is the share of employment in the construction industry in the years prior to the bust. All else equal, the analysis examines whether these three factors had a measurable and negative impact upon house price growth in 2005 through 2010. All of these preconditions need to be addressed in order to identify the underlying structure in the current market and the one that will ultimately emerge when the effects of these preconditions finally dissipate.

The third challenge pertains to largely unanticipated repercussions during the Great Recession. The dominant example of this is the widespread incidence of foreclosures as well as properties with substantial negative equity, both of which have led to the growth in what is often referred as the inventory of distressed housing. This inventory includes vacant properties and others in various stages of the foreclosure process or as potential candidates for foreclosure. This inventory can cast a severe shadow over normal market operations and increase the difficulty of identifying the long-run or emerging structure of the market housing market. For example, the estimated relationship between the level of house prices, income, rent, and other normal drivers of the reduced form equation may be quite different in markets with a small inventory of distressed real estate versus one with a much larger inventory. Our approach utilizes data on distressed real estate developed by Collateral Analytics and provided for this study. A variety of measures of the distressed real estate inventory are available. This paper focuses upon the percent of the stock of owner-occupied housing with a mortgage that has been foreclosed and is awaiting sale back to the private market. This measure is often referred to as the stock of Real

[1] A high cost mortgage for this period is one in which the APR is at least three percentage points above rate on a Treasury bond of similar maturity to the mortgage. See http://www.ffiec.gov/hmda/glossary.htm#R.

Estate Owned (REO) properties because these properties are owned by financial institutions and servicers who are anxious to sell them on the private market to recoup some of the losses associated with the loan on the property.

Why is it important to identify the underlying structural relationships that drive house prices? Here are several examples.

- First, the experience of this Great Recession provides a potentially valuable opportunity to measure the substantial secondary effects associated with the bubble and bust. These were unanticipated in the previous models and in most other models of house prices in the literature developed prior to the current bust. Future models of house prices may want to incorporate these ancillary impacts in order to provide policymakers and investors with a fuller appreciation of the costs of a bubble bust.
- Second, the results may also provide further impetus or motivation for alternative policies to help bring about a return to normalcy in housing markets. Examples of policies of this type include macroprudential policies launched by the Federal Reserve that are designed to curb potential bubbles and bust such as cyclically oriented capital standards for financial institutions. They also include "microprudential" policies that facilitate the conversion of REO properties into rental properties.
- Third, the analysis may identify places that are particularly hard hit by the crisis and where prices are far below their equilibrium values and where some type of public assistance or policy intervention may be most in need or justified. Indeed, all places were not hit the same. As such, policies designed for specific local markets may be required. The results of this study may help identify these places and highlight the wide variability of the current housing market.
- Fourth, the results may shed light on questions such as these: where are housing prices headed and what does the new normal look like? Perhaps it will be possible to identify certain new measures of the housing market, especially in those places that have been hit less hard or that have recovered most quickly. It may also identify areas in which the road to recovery is particularly slow and long.

These are all examples of policies that rely upon the type of results generated in this paper.

Section II of the paper reviews some of the massive literature on the current housing price bubble and bust to motivate the selection of the variables focused upon in this paper and to place the contributions of this paper in the context of the existing literature. Section III discusses the very large data set that has been assembled with a focus on the measures of distressed real estate used, the core list of exogenous variables, and some of the challenges related to the selection of the best house price index. Section Four discusses the modeling approach and the criteria to be used in the search for structure. Section Five presents the main findings. The final section summarizes key conclusions, a brief discussion of the relevance of work of this type to the current policy debate about how to facilitate a return to "normal." An appendix contains summary statistics for many of the variables used in the paper.

Section II: Literature Survey

The purpose of this section is to use some of the massive and still growing literature on the housing market to highlight some of the challenges that affect the search for structure in today's housing markets. It begins with a brief look at the traditional models of house prices, which rests upon a detailed structural model of the housing market and then moves to a discussion of a modern descendant of the traditional structural model that dominates much of the literature today: the VEC approach. A third branch of the literature is then discussed; this branch is labeled in this paper as "Detailed Case Studies" and refers to empirical studies that focus upon a relatively narrow dimension of a structural model of the housing market such as the rent to value ratio. The final part of the literature review includes a discussion of some of the specific challenges faced in a search for a stable structural picture of the housing market during the current crisis. In each part, examples from the literature are used to help motivate our taxonomy of the literature. The final part closes with an attempt to position the contributions of this paper within the literature.

Traditional Literature: A Detailed Structural Model of Housing Supply and Demand

The roots of any model of the housing market lie in its characterization of the key elements of the structure of the housing market. Such a model focuses upon various aspects of the core drivers of the supply and demand for housing. The demand equation posits that demand is a function of the price of housing and household income. The supply equation addresses the aggregate quantity of housing as a function of the price of housing and key drivers of the cost of providing the housing, e.g. construction costs. Estimates of the parameters of these models usually focus on the price elasticity of the demand, the income elasticity of demand, and the price elasticity of supply. Inferences about the effects of certain market or policy shocks are obtained by linking the shock to one of the core exogenous variables in the system and then simply computing the net effect on the price and quantity of housing with the various elasticity estimates. Relative to the modeling of housing that was to emerge in the 1990s, this approach featured a relatively strong reliance upon a theoretical model of the housing market and reflected a world in which the amount of data about the housing market was quite limited. Simultaneous equations estimators like three-stage least squares are typically used to estimate the structural equations; the reduced form models could be estimated by regression.

An example of this approach is the model estimated by Blackley and Follain (1991). They posit a five equation model. Two pertain to the market for owner-occupied housing. Two pertain to the market for rental housing. A fifth equation examines the drivers of the home ownership rate. The model is estimated using two years of cross-section data for thirty-four MSAs from the late 1970s. This paper and many others confirmed the strong role of household income, a substantial own price elasticity of demand, and a generally elastic supply curve for housing. A great challenge in this approach is the need to develop measures of the quantity and price of housing. Hedonic price indexes play a major role in such work. Also, these models have difficulty capturing the dynamic properties of housing markets, the interdependence of housing with other sectors, and the strong role that local market conditions can play. Incorporating the role of mortgage debt, mortgage choices are other examples of the challenges facing the detailed structural model approach.

VAR-VEC Approach

Many of the challenges faced by the traditional structural approach became increasingly clear as the housing market evolved and much more data became available to study these markets. These challenges were also being faced in the field of macroeconomics where attention was focused on other markets such as the supply and demand for labor and investment. Many macroeconomists responded to these challenges by shifting to a VAR-VEC approach, which might be considered a "descendant" of the traditional structural model. These were designed to capture the dynamics of change over time and the interaction among the endogenous and exogenous variables in the structural model. Theory guided the selection of variables in the system, especially in the first stage VEC equation, but the final conclusions and insights were greatly influenced by what the data and the estimates had to say. This approach has become very important in the search to explain house price bubbles. The first equation in these models—the VEC—seeks to capture the reduced form for house prices. The other equations in the model—the VAR equations—focus upon growth rates in house prices and their potential interaction with some of the drivers of house prices. Deviations of the actual level of prices from the level predicted by the VEC equation also played a role in predicting future house price growth. These models do not yield a simple picture of the various elasticities that readily follow from the structure models. Rather they are inferred from various types of shocks or simulations of the model parameters via impulse response analysis. Estimates of location specific fixed effects are also typically included to capture the myriad of local market conditions that affect housing outcomes. Abraham and Hendershott (1996), Malpezzi (1999) and Dreiman (2004) provide examples of this approach prior to the analyses of the current bubble-bust.

Detailed Case Studies: Deep Dives and Looking Under the Hood

Another category of the literature on housing markets can be dubbed Detailed Case Studies. These seek to examine specific issues or market traits that influence the level of house prices. In some sense, these seek to take a closer look "under the hood" to see what factors drive housing markets that are not typically reflected in either the structural or the VAR/VEC models.

Here is a very brief summary of some of this type of approach.

- <u>Declining Cities and the Asymmetric Supply Curve</u>: Glaeser and Gyourko (2005) examine differences in a reduced form equation for housing between growing and declining cities. The idea is that the supply curve is much less elastic in a declining city where the outmigration of people and jobs has produced an excess supply of housing. Follain (2011) also examines this issue. Both show support for this asymmetric supply curve.
- <u>Regulatory indexes</u>: The supply elasticity for a market reflects local regulatory policies. Gyourko, Saiz, and Summers (2008) develop and analyze a new survey of over 2000 jurisdictions across all major housing markets in the US and document how regulation of residential building varies across space. Coastal markets tend to be more highly regulated, with communities in the Northeast region of America being the most highly regulated on average, followed by those in the West region (California especially).

- Physical influences on Supply: Saiz (2010) processes satellite-generated data on terrain elevation and presence of water bodies to precisely estimate the amount of developable land in US metro areas and show that geography is a key factor in the contemporaneous urban development of the United States. The data show that residential development is effectively curtailed by the presence of steep-sloped terrain and that most areas in which housing supply is regarded as inelastic are severely land-constrained by their geography. Econometrically, supply elasticities can be well characterized as functions of both physical and regulatory constraints, which in turn are endogenous to prices and demographic growth.
- Interaction of Supply Elasticity and Price Volatility: Glaeser, Edward L., Gyourko, Joseph, and Saiz, Albert (2008) conduct an empirical study that incorporates housing supply into the study of house price movements. They present evidence that shows that the welfare consequences of surges in house prices may actually be higher in more elastic places because those places will likely experience more overbuilding in response to a surge in house prices. They show that the price run-ups of the 1980s were almost exclusively experienced in cities where housing supply is more inelastic. More elastic places also had slightly larger increases in building during that period.
- Spatial statistics: Lu and Wang (2010) investigate and demonstrate a strong spatial connection among housing markets. Previous research argued that spillover effects of an economic shock on the housing market would not necessarily occur between neighboring areas, but would be based on economic interrelationships. This paper uses spatial panel models to investigate the spillover mechanism in the housing market. The empirical regression results showed that both the inter-MSA spatial correlations in housing returns are significant and positive, but with the geographic correlation appearing to be stronger and more significant than the correlation based on economic similarities. Their empirical investigation also provided evidence that housing price changes in one MSA area are correlated not only to their own lagged changes, but are also correlated to the present and lagged housing prices changes in their neighboring or economically similar MSAs.
- Interaction among Markets: Miller, Peng, and Sklarz (Forthcoming) and Zhou (2010) use quarterly data for all 379 metropolitan statistic areas (MSAs) in the U.S. from 1980:1 to 2008:2 to study, the effect of house prices on local Gross Metropolitan Product (GMP). They compare the effects of predictable and unpredictable house price changes, to capture the collateral and wealth effects of house prices respectively. Their analysis provides the following findings. First, house price changes have significant effects on GMP growth, and the effect of predictable changes (the collateral effect) is about three times stronger than the effect of unpredictable changes (the wealth effect). Second, the persistent component of predictable changes has a stronger collateral effect than the novel component. Third, when households are more financially constrained, the collateral effect is stronger, the wealth effect is weaker, and the total effect remains unchanged. Finally, the effects last for eight quarters, and peak on the fourth quarter after house price changes take place. Zhou (2010) also investigates the interaction between housing and the broader economy.

Now attention to the huge literature focused upon the current bubble and bust.

What Is Different this Time?

The depth of the declines in house prices during the current bubble bust have prompted an extraordinary response from the economics profession to explain what happened. The literature has been buttressed by vast amounts of data not previously available and by a wide variety of hypotheses put forth to explain what happened. A discussion of this segment of the literature begins with a brief summary of two papers that offer high level overviews of the literature about the crisis. This is followed by a discussion of some papers within the literature that are closely related to the approach taken in this paper.

High Level Reviews. Two papers published in 2012 offer interesting and high level reviews of the various explanations being offered to explain the origins of the current crisis. A number of themes emerge from these studies that this paper hopes to emulate. One is the difficulty of uncovering the "truth" associated with this crisis and the need for economists to be humble about the judgments offered about the causes of the crisis.

One is by Foote, Gerardi, and Willen (2012). This paper presents 12 facts about the mortgage market. The authors argue that the facts refute the popular story that the crisis resulted from financial industry insiders deceiving uninformed mortgage borrowers and investors. Instead, they argue that borrowers and investors made decisions that were rational and logical given their *ex post* overly optimistic beliefs about house prices. The authors then show that neither institutional features of the mortgage market nor financial innovations are any more likely to explain those distorted beliefs than they are to explain the Dutch tulip bubble 400 years ago. They encourage economists to acknowledge the limits of our understanding of asset price bubbles and to undertake the design policies with this in mind. They say that the study of bubbles is too young to provide much guidance about why this bubble occurred. "For now, we have no choice but to plead ignorance, and we believe that honest economists should do the same."

The other is by Lo (2012). The recent financial crisis has generated many distinct perspectives from various quarters. In this article, Lo reviews a diverse set of 21 books on the crisis, 11 written by academics, and 10 written by journalists and one former Treasury Secretary. No single narrative emerges from this broad and often contradictory collection of interpretations, but the sheer variety of conclusions is informative, and underscores the desperate need for the economics profession to establish a single set of facts from which more accurate inferences and narratives can be constructed. Lo offers several observations that stem from his review of these various narratives. "The most obvious is that there is still significant disagreement as to what the underlying causes of the crisis were, and even less agreement as to what to do about it. But what may be more disconcerting for most economists is the fact that we can't even agree on all the facts."

Examples of the VAR-VEC Approach. Follain and Giertz (2011, 2012) provide an example of this approach and a discussion of the large literature in which the VEC-VAR model is used to study house price bubbles. The first of these provides estimates of model using the MSA as the unit of observation for various years between 1980 and 2010. The level of real house prices in the first stage is driven by income per capita, employment, interest rates, and various fixed effects. The residuals from this equation are incorporated into the 3 equation VAR model designed to explain the growth in house prices and the interconnectedness among the three main variables. Impulse response analysis is used to make inferences about the impacts of various shocks on house prices. A key result of this paper and many others in the field is the strong role played by momentum, which is generally not incorporated into the traditional structural models. A specific result of their work is the wide range of predictions about future house prices and the difficulty of predicting those markets that were particularly hard hit by the recent crisis. The second paper uses quarterly data from 1990 through 2011 and considers other potential drivers of house prices such as the volume of sales and the level of residential rents. A key finding of this paper is that the model forecasts can change rather suddenly during the final stages of the boom. Both papers also provide detailed reviews of many other papers that use this approach to study the current bubble and bust, e.g. Lai and Van Order (2010).

A key debate in this literature pertains to the specification of the reduced form equation for housing in the first stage VEC. The goal is to incorporate the relevant variables that affect the equilibrium price of housing. Departures from this equilibrium price are indicators of the potential of a bubble or bust. Here is a brief summary of some of the ways in which the equilibrium price is depicted in this literature.

- Rent to value ratio: Leamer (2003) championed the potential role of the rent to value ratio as a signal of an upcoming bubble. The intuition underlying this variable is powerful and can be motivated from a variety of fields besides housing. Lower rent to value ratios, all else equal, imply that investors are expecting more appreciation. Campbell, Davis, Gallin, and Martin (2009) examine the stability of the relationship with MSA data through 2007 and highlight the importance of controlling for variations in the variability of the risk premium for housing and the interaction among components of the rent to value ratio over time. Goodman and Thibodeau (2008) and Follain and Follain (2007) also focus on this variable. The challenge that has become more clear with time is that all else may not stay equal during turbulent times and the assumption of a stable rent to value ratio, especially during turbulent times, is a strong one in need of confirmation
- Price to Income ratio: In their 2004 paper entitled Is There a Bubble in the Housing Market?, Case and Shiller focus upon the ratio of price to income per capita as an explanation of house price growth in years prior to the current bubble—1985–2002. They find a strong connection in most states. Indeed, income alone almost completely explains home price increases in the vast majority of states, about eight states are characterized by large swings in home prices that exhibit strong inertia and cannot be well explained by income patterns. These included six states in the Northeast, California, and Hawaii, which may be considered similar to the SAND states (Arizona, California, Florida, and Nevada) of the current crisis. Dreiman-Harter (2000) wrote a short but prescient piece on the FHFA web site in which she highlighted the interaction between the price to income ratio and demographic factors. She pointed out that the larger the share of the population comprised of younger households in their 20s, the higher the price to income ratio. In 2000, the price to income ratio was relatively low and some considered this a sign that a bubble was not on the horizon, Dreiman-Harter suggested that the historically low ratio may have been reflecting an aging population in which the share of the population in their 20s was five percentage points lower than in the late 1970s.

- Sales volume: Miller, Peng and Sklarz (2011) empirically analyze the effects of single family home sales on house prices and the local production (GDP). They use data for 372 Metropolitan Statistic Areas in the U.S. from the first quarter of 1981 to the second quarter of 2008 in a panel Vector Error Correction Model. Changes in home sales are found to Granger cause the growth rate of per capita Gross Metropolitan Product, and the dynamic effects are visualized with impulse response functions. Supporting evidence for the economic impact of home sales on house prices is also found in contemporaneous regressions in their 2011 paper.
- Applying parameters from a stable period: Another approach to bubble detection involves using estimates of a model estimated during a period of stability to one suspected of being in the bubble formation. This, in essence, seeks to incorporate all of the potential indicators of a bubble into a single or multi-equation model. If the model from a stable period understates the level of house prices, all else equal, then potential evidence of an emerging bubble is said to exist. Goodman and Thibodeau (2008) use this approach as well. Wheaton and Nechayev (2008) also use this approach and relied upon models estimated through the mid-1990s as the basis of their assessment of whether a bubble existed in the early 2000s. They examine whether the inflation in housing prices between 1998 and 2005 can be explained by increases in demand fundamentals such as population, income growth and the decline in interest rates over this period all of which are included in their model. In all of the 59 MSA markets in their study, the growth in fundamentals from 1998–2005 forecasts price growth that is far below than which actually occurred. They also examine the magnitude of the 2005 forecast errors in a cross-section regression. This regression found that the errors were greater in larger MSAs, in MSAs where 2nd home and speculative buying was prevalent, and in MSAs where indicators suggest the sub-prime mortgage market was most active. They also conclude with an important comment: because the subprime mortgage developments were unique to the housing market of the early 2000s, it was difficult to assess if and how far housing prices would "correct" after 2005.

Case and Shiller (2004) also offer a nice overview of these various indicators of a bubble and conclude that they all are less than ideal predictors of emerging bubbles. They refer to household surveys that they have conducted in several markets for many years as an alternative way of learning about the potential of excessively optimistic expectations of future house prices. They conclude that these surveys may be a better indicator of what is to come than are models built upon these various indicators of market equilibrium. In so doing, they underscore a challenge facing any econometric model of house prices—keeping track of heterogeneous and evolving household expectations influenced by a variety of information not easily incorporated into econometric models.

Examples of the Many Detailed Case Studies Spawned by the Crisis

The current attempts to understand the causes of the recent bubble include a number that fit under the category I defined above—detailed studies that explore one or more elements of a comprehensive structural model. Here are some examples of this genre of the literature

Agarwal et al (2012) examine another of the externalities associated with the last bubble bust. This paper focuses on the potential externalities associated with subprime mortgage origination activity. Specifically, they examine whether negative spillover effects from subprime mortgage originations result in higher default rates in the surrounding area. The empirical analysis controls for loan characteristics, house price changes, and alternative loan products. The results indicate that after controlling for these characteristics, the concentration of subprime lending in a neighborhood does not lead to greater default risks for surrounding borrowers. However, they do find that more aggressive mortgage products (such as hybrid-ARMs and low/no documentation loans) had significant negative spillovers on other borrowers. Stated differently, the aggressive alternative mortgage designs and their ability to generate more highly leveraged housing investments were more toxic to the housing and mortgage market than previously believed.

Barlevy and Fisher (2010) describe a rational expectations model in which speculative bubbles in house prices can emerge. When a bubble emerges, both speculators and their lenders prefer interest-only (IO) mortgages to traditional mortgages. By contrast, absent a bubble there is no scope for mutual gains from using IOs. Using data compiled for over 200 US cities for the period 2000–2008, they find that IOs were used sparingly in cities where elastic housing supply kept house prices in check, but were common in cities with inelastic supply where house prices rose sharply and then crashed. They confirm that the use of IOs in these cities is not proxying for other mortgage market characteristics such as subprime, securitization, or high leverage. Moreover, the use of IOs does not appear to have been a response to houses becoming more expensive; if anything, their use anticipated future appreciation. They also confirm that IOs were more likely to be repaid early, and those that survived until prices fell were more likely to default. These findings suggest that the recent boom-bust in the housing market was associated with a speculative bubble.

Archer and Smith (2011) examine the influence of exotic mortgage loans in the house price bubble using data from 20 counties in Florida. They find evidence of their impact on the bubble, but their panel VAR impulse response analysis also confirms a more important role for increased momentum as a driver of house prices in the early to mid-2000s.

Agarwal, Ambrose and Yildirim (2010) develop a theoretical model demonstrating the potential spillover effects associated with the introduction of risky assets. Specifically, they examine the potential increase in mortgage default risk on prime mortgages that results from the introduction of subprime mortgages in a local area. The transmission mechanism is that the higher incidence of default by subprime borrowers leads to greater asset price volatility, which in turn impacts the default probability of prime mortgages in the same geographic area. Through numerical analysis, they demonstrate the impact of the origination of subprime mortgages to the risk of a portfolio of prime mortgages. Finally, they offer empirical support for their model by examining the spatial variation in MSA prime mortgage default rates correlated with the level of subprime mortgage activity.

Chen, CarBaCho-Burgos, Mehra, and Zoller (2011) describe the model of house prices developed by Moody's Analytics. Their approach is based on a structural model of housing demand and supply that allows for serial correlation and mean reversion in metropolitan area housing markets. The model that Moody's Analytics has developed is a tool for identifying the forces driving house prices and assessing to what degree house prices can be explained by fundamental, persistent forces and to what degree they are explained by more temporal forces. They pay special attention to the interplay between house prices and foreclosure activity and find a strong and negative impact of foreclosure sales on house prices.

A number of other papers examine some of the institutional context that affected house price outcomes and the externalities associated with them. Morgan, Iverson, and Botsch (forthcoming) examine the connection between bankruptcy laws and subprime defaults. Campbell, Giglio, and Pathak (2009) use data on house transactions in the state of Massachusetts over the last 20 years to show that houses sold after foreclosure, or close in time to the death or bankruptcy of at least one seller, are sold at lower prices than other houses. Landvoigt, Monika, and Schneider (2012) use data at the zip code level in San Diego to study how equilibrium house prices adjust to assign houses that differ by quality to movers who differ by age, income and wealth. Their main result is that cheaper credit for poor households was a major driver of prices, especially at the low end of the market. Mian, Sufi, and Trebbie (2011), Kiefer and Kiefer (2009), and Follain, Miller, and Sklarz (2012a and 2012c) are other examples of work designed to capture measure the interplay between house prices and the various stages of the foreclosure inventory.

The final paper discussed in this review of recent attempts to explain the ongoing housing crisis is by Glaeser, Gottlieb, and Gyourko (2010). Their results are interesting, but they also highlight the challenge of untangling the interplay between house prices and mortgage lending in the years leading up to the bubble bust. They focus upon one of the explanations of the boom: access to easy credit in the form of low real interest rates, high loan-to-value levels and permissive mortgage approvals. They use the standard user cost model of housing prices and conclude that the predicted impact of interest rates on prices is much lower once the model is generalized to include mean-reverting interest rates, mobility, prepayment, elastic housing supply, and credit-constrained home buyers. The modest predicted impact of interest rates on prices is in line with empirical estimates, and it suggests that lower real rates can explain only one-fifth of the rise in prices from 1996 to 2006. Thus they find no convincing evidence that changes in approval rates or loan-to-value levels can explain the bulk of the changes in house prices, but acknowledge that definitive judgments on those mechanisms cannot be made without better corrections for the endogeneity of borrowers' decisions to apply for mortgages.

Potential Contributions of this Paper: Uncover Emerging Structure

The specific steps undertaken in this study build upon the VAR/VEC approach. The focus is upon the years during the crisis—2005 through 2010. This necessarily limits the length of the time-series. On the other hand, the use of county data offers a more geographically granular approach than many of the studies. These more recent data also provide an opportunity to examine a wider variety of price measures only available in recent years

Of particular interest is whether a stable reduced form equation can be found and whether residuals from it play an important role in explaining house price change during the crisis. The approach also seeks to distinguish three sets of factors. One consists of the "usual suspects" that typically define a reduced form or equilibrium level of house prices. Another seeks to capture the conditions prior to the bust such as the prevalence of subprime or high cost lending, the growth in house prices prior to the bust, and the share of employment in the construction industry prior to the bust. A third set seeks to incorporate the interplay between the growth in foreclosure activity and house prices. The analysis also seeks to follow the advice of both Foote et al and Lo—do not overstate the significance of our findings and try to leave no "reasonable stone unturned."

Section III: Data

The development of the data for this paper sought to accomplish four goals. One is to expand the list of potential explanatory variables for the recent bubble-bust relative to those available for the two previous papers by Follain and Giertz and in many other papers about the recent bubble-bust. A second goal is to look at the housing market at the county level versus the MSA level used in the previous papers, which requires additional house price measures since a publicly available index of house prices is not available at the county level. A third goal is to incorporate measures of the distressed real estate inventory, which is an area in which much work and many measurement options are being pursued by a wide variety of agents. Publicly available data regarding these patterns are quite limited. A final goal is to focus upon the period 2005–2010, which encompasses the end of the boom and the ongoing bust.

All of these considerations led to a major effort to assemble a large data set with the requisite information and a wide variety of variables of potential drivers of house prices at the county level. The data set is comprised of four major sources: American Community Survey; Collateral Analytics; Moody's Analytics; and a variety of miscellaneous sources. Each is summarized briefly in this section.

American Community Survey (ACS) Data. The ACS data are available for 638 counties on an annual basis from 2005 through 2010.[2] Data for 2011 are not yet available. The process consists of the identification and downloading of a large number of variables from the web site at the county level. The focus has been upon variables regarding housing, employment, population, and income. Each variable comes with a county and year level estimate and an estimate of the confidence interval for each variable. One especially important variable is built upon the estimates of house value by the owner-occupants of these housing units. The median value of the owner estimates provides an index of the value of owner-occupied housing units among counties and over time. Several other key variables from the ACS include median gross rent, median household income, the number of owner-occupied housing units and the number of these units with at least one mortgage.

Collateral Analytics Data. Collateral Analytics (https://collateralanalytics.com/) is a private firm devoted to the measurement of the value of the housing stock and a variety of products designed to measure and forecast trends in house prices. The major source of these data is public record data from all over the country, which typically comes from assessor records and is supplemented by a variety of sources. Collateral Analytics (CA) has provided a number of critical variables for this study. These include an index of the price per square foot, the number of sales, the number

[2] Barbara Follain played an important role in assembling the American Community Survey from the census web site: http://www.census.gov/acs/www/data_documentation/data_main/.

of REO sales, the number of foreclosures, and the size of the inventory of properties on which foreclosure has occurred and the property awaits sale back into the private sector by the banks that own the properties (REO sales). Follain regularly uses these data in ongoing writing projects with Michael Sklarz, the President and founder of Collateral Analytics. Two recent examples include short articles about the distressed real estate inventory by Follain, Miller, and Sklarz (2012a and 2012c), which discuss the various stages of the distressed real estate inventory and estimates of the size of this inventory for four counties in New York and four counties in California.[3]

Moody's Analytics. Moody's Analytics is a major consulting firm well known for the data it collects and the research it conducts with these data. The data are made available to the public at its web site: http://www.economy.com/default.asp. These data are usually purchased, but Moody's has provided us with information for 20 or so variables at the county level for this study.[4] The primary sources for these data are typically government based but Moody's adds great value by assembling these data. The main variables selected for this study were meant to complement and sometimes overlap with the ACS and Collateral Analytics variables.

Miscellaneous Sources. These include the Federal Housing Finance Agency (FHFA) house price index, fair market rents from HUD, and various measures of the physical location and traits of counties. Price index data on rental housing were downloaded from the web site of Edgar Olsen http://eoolsen.weebly.com/price-indices.html. Some additional data regarding the Wharton Regulatory Index and estimates of MSA supply elasticities were obtained from the web site of Albert Saiz (http://real.wharton.upenn.edu/~saiz/).

The data assembled for the project now consists of over 400 variables. Many of these are transformations of the raw variables, which number at least 200. The appendix contains basic statistics on a large number of the variables employed in this paper. Additional explanations of the main variables are offered below and in Section V.

One challenge is to assess the quality of the information available from these sources regarding the value of housing. Each has its advantages and potential disadvantages. What follows is a review of the thinking underlying our decision to focus upon a price index built upon Collateral Analytics data. Four sources of information are available about the level of prices at the county level. These include the ACS median owner estimate of value (HP_ACS), the median price per square foot of the residential regular and REO property sales (HP_CA)); and two measures of the median sales price, one from CA (HP_CA2) and one from Moody's (MSP). A fifth variable, the FHFA price index, provides information about changes in prices over time at the MSA level. This is assigned to each county in the MSA with an FHFA index.

Some judgments were made to compile these data and the ultimate number of counties used in the estimation. One challenge is that the public records data are not available for all of the ACS counties for each year. In those cases in which the CA data were missing or based upon very

---

[3] I am grateful to Mike Sklarz for providing these data for this study and the work by his staff to provide the data and help understand it better.
[4] I am grateful to Mark Zandi and Celia Chen of Moody's Analytics for agreeing to provide the data for the purpose of this study.

small sample sizes, data from either Moody's (MSP) or the new CA median sales price series (HP_CA2) were used. An alternative price index is also constructed by multiplying the adjusted price per square foot series from CA data by 1,750 square feet to provide an index of this constant size house over time and among counties. This series provides a close match between the average ACS value based upon the median value of owner estimates of value and the average of HP_CA. The number of counties with requisite data in the ACS, Moody's, and Collateral Analytics data base is 439 for 2006–2010 and 416 for 2005. Table III-1 offers the average values of these indexes for each year in the sample.

The average value of the CA based price index (HP_CA) is $244,654 in 2005 and has declined to $231,248 in 2010 (these are averages of quarterly data). The fifth column contains information for HP_CA in the 416 counties in which these data are available for each year of the sample. This is relevant to estimation procedures that rest upon a balanced time-series and cross-section panel. The averages among the five data series are similar, though they vary substantially among counties and over time. The various rows of Table III-1 capture some of the changes in these variables over time. The cumulative percentage changes in each of these variables are provided in Table III-2. These highlight a key factoid of the data: the HP_CA and the MSP series from Moody's generally show larger declines than the other series.

Table III-3 focuses upon the index that will be used for most subsequent analysis in this paper: the HP_CA series. These data highlight variations in the price indexes for different groups of counties. The first column includes all counties by year and the others focus upon SAND (Arizona, California, Florida, and Nevada) versus nonSAND states and three different groups of counties based upon their longer term population growth (slow (Dpop=1), medium (Dpop=2), and fast (Dpop=3). Among all counties, the peak in prices occurred in 2006 and since then prices have declined by an average of 20 percent. They are 17 percent below their 2005 average. The comparable declines among the counties in the SAND states were much larger, 41 percent from the peak of 2005 and 38 percent since 2005, than among the nonSAND states. The slow growing counties experienced lower average declines than the other two groups, which were pretty similar on average. The indexes are plotted in Figure III-1 for various subgroups.

One key judgment underlies the choice of the HP_CA index as the primary focus of attention in the estimation to follow. It rests upon a pro and a con of the ACS index relative to the CA index. On the one hand, the ACS variable refers to all owner-occupied housing whereas the CA index focuses upon properties that have sold. This is a positive aspect of the ACS data, all else equal. On the other hand, the ACS estimate is based upon the owner's estimate of the value of the property, which may be biased.[5]

One specific concern is whether owner estimates are in tune with the turmoil in the market. Follain (2009) discusses the possibility of "Stubborn Sellers," who are households who may have based their estimate of value for the price they paid in the early 2000s and who are now reluctant to accept what the market may be offering during the crisis. If so, homeowner estimates of value may be upwardly biased. This issue is examined by generating a measure of the gap between the ACS value and the HP_CA value (HVALgap).

---

[5] There is a large literature on this topic, which continues to this day.

A tentative judgment based upon a review of the data and some simple tabulations and regressions is that the bias may be especially substantial in the areas hit hardest by the crisis. The tabulations in Table III-4 show that the owner estimates are consistently higher than the HP_CA values in the SAND states in recent years. They were actually underestimating values by about 13 percent, on average, during the last year of the boom in 2005. Now owners in these areas are overestimating by 18 percent in 2010 compared to overestimates of 8 percent in the nonSAND states. Another insight on this point comes by listing the 27 counties with the largest HVALgap in 2010, which is done in Figure III-2. The value of HVALgap exceeds 25 percent in each of these counties. The MSA in which the county is located is plotted on the horizontal axis and the size of the gap (HVALgap) is on the vertical axis. All of the counties represented in this group are in the SAND states. Three counties are in the Jacksonville, FL MSA (St. Johns, Clay, and Duval). Lee County in the Cape Coral, FL MSA has the largest gap at 64 percent.

A last item of attention in this section regarding the data relates to the level of rents and two ratios often used as indicators of potential bubbles: the rent to value ratio and the price to income ratio. Rent information is available from the ACS and HUD's Fair Market Rent data. These data provide a variety of ways to measure rent and the two ratios. Some of this information is summarized in Table III-5. Note that the unadjusted mean of the ACS data is quite similar to that for HUD's 2 bedroom rental unit. In 2005, these are $8,999 and $9,145, respectively; in 2010, these are $10,506 and $10,866. The ACS values are deflated by a rental price index series generated by Edgar Olsen and colleagues. This variable is called the ACS-Rent_Olsen Adjusted in the table and is the one focused upon in this study.

This measure of the cost of quality-adjusted rental housing is divided by the HP_CA price index data to generate rent to value ratios (Olsen_Adj Rent to HP_CA). As discussed in the previous sections, the R/V ratio is often used as a potential indicator of a potential bubble. A very low ratio, all else equal, is considered to be an indicator of higher expectations of future price growth. More on this issue is presented below. Here simply note the average value of this ratio and the wide variation in it. The mean in 2005 was 4.3 and 5.2 in 2010. The standard deviation is 10 percent higher in 2010 and is as high as 14.3 percent in one county. The two counties with the highest ratios in 2010 are in Texas and have ratios above 10 percent. The last item of note in Table III-5 is the price to income ratio, which is based upon the ratio of HP_CA relative to the ACS median household income. The average of this ratio is 4.6 in 2005 and 3.7 in 2010. The dispersion seems to have diminished since 2005. Results are presented later in the paper to ascertain the stability of these two ratios among counties and during the distressed environment.

Section IV: Estimation Strategy and Search Criteria

The ultimate goal of this paper is to identify what may be considered a reasonable and evidence based specification of key parts of a structural model underlying house prices. The focus is upon the period 2005 through 2010, which is one of substantial turmoil. This section discusses the equations on which the search focuses, the estimation techniques, and the criteria used to assess whether the search for the underlying structure has been successful.

Three sets of equations are estimated.

- The first set includes reduced form equations for three variables of interest: the level of house prices; the rent to value ratio (or rent to price ratio); and the value to income (or price to income ratio). The explanatory variables are similar for each of these equations and include a mixture of variables thought to be exogenous. Some experimenting was done to select and define this list of variables for inclusion in the models. The final list includes some of the traditional drivers of house prices, some that seem reasonable and perform well in the model, and some that seek to capture the level of distress in the markets and certain preconditions that have had an enduring impact on the level of house prices. The equation for the level of house prices can be considered the first stage VEC equation.
- The second set of equations focus on estimates of the growth rates in the level of house prices. A VAR specification is used, which means that the right hand side variables include the residual from the VEC, lagged growth in house prices, and several key variables that interact with house prices.
- The third set of equations focuses upon the set of county fixed effect estimates generated by one of the estimation processes used to estimate the VEC model

The discussion of the estimates of each set of equations focuses upon the particular variables in the equations and their stability among counties and during the 2005–2010 time period.

Three types of estimation techniques are used. One is standard regression analysis with a variety of right-hand side variables that include the traditional drivers of house prices and various ways to depict county specific fixed effects. A second approach uses forward looking stepwise regression to help select regressors from a long list of potential candidates. The third approach is a specific procedure especially designed for panel data sets of the type used in the analysis. The procedure in Stata to do this estimation is called xtreg.[6] It seeks to identify county specific factors that affect the dependent variables. Each output of this estimation procedure generates estimates of these fixed effects for each county in the panel. They are essentially equal to the average of the residuals for each county over the six year period. These fixed effect estimates become the dependent variables in the third set of equations to be estimated. The equations to explain the county fixed effects are estimated for all of the counties and some subgroups of the counties.

Several criteria are used to populate the various equations to be estimated. First, the regressors in any of the equations, but especially in the reduced form equations (VEC), are considered to be exogenous and can be justified as a determinant of the demand or supply of housing in a county. Second, the residuals from the first stage are reviewed to identify places where the model fits particularly well and not well and also incorporated into the VAR growth rate equations. Third, the growth rate equation includes variables thought to be strongly related to house prices and possibly jointly influenced. Fourth, a list of possible candidates to explain the fixed effects is

[6] These can be modeled as random or fixed effects. The results in this paper utilized the fixed effect (fe) option in Stata. More information about this can be found at: http://www.stata.com/capabilities/cross-sectional-time-series-panel-data/.

drawn from the literature about the supply elasticity of housing, the regulatory environment that affects the supply, and various physical descriptions of the county landscape. Lastly, the predictions of the growth rate equation for 2011 are evaluated to identify places where the predictions were very good or very poor. The assessment of the search process stems from comparisons of the results for each set of equations to what is expected from previous research and by the stability of the estimates among counties and time periods. Substantial deviations of the estimates from previous research and widely varying estimates among counties and over time are judged to be evidence of an unstable or evolving structure. The predictive power of the estimated equations is also used as a barometer of potential success; larger R^2 values, all else equal, are suggestive of a stable structure, all else equal.

In sum, a mixture of drivers of house prices is considered; some are theoretically plausible variables and some are based upon what the results and data suggest as other strong empirical regularities. Various estimators are used and sometimes, as is noted below, they yield quite different coefficient estimates for the same variables. Judgments are made to label our preferred structural models. As noted at the end of the literature survey, all of this is intended to be done and presented with the humble attitude suggested by Lo (2012) and by Foote, Gerardi, Willen (2012) and many others like Taleb (2007). An effort is made to avoid or at least be cognizant of judgments underlying the conclusions of the paper that are based upon weak empirical evidence and that could lead to very poor predictions if the judgments prove in the future to be poor.

Section V: Key Results

The analysis of the estimation results focuses upon a search for stable relationships within the data. This begins with an exploration of a reduced form equation for the level of house prices (HP_CA). Four other equations are also estimated:

- an equation for growth rates in house prices;
- an equation for the fixed effects from the reduced form equation for the level of house prices using HP_CA;
- an equation for the rent to value ratio; and
- an equation for the price to income ratio.

The exploration considers various specifications of the right-hand side variables, the variability of the estimated relationships among subgroups of the counties, and a review of their residuals, which helps to identify those areas in which the estimated relationships fit particularly well or poorly.

Search for a Stable Reduced Form Model

The dependent variable in the reduced form model is the natural log of the level of prices using the Collateral Analytics index of house prices (lnHVAL_CA). The reduced form equation is important in its own right because it expresses the main drivers of the level of house prices. A vector of the residuals from the reduced form is also used as an explanatory variable in the estimation of an equation to explain the growth rates in house prices.

A variety of right-hand side variables for the reduced form is considered. Some of these variables are strongly suggested by theory such median household income (lnincome_ACS), the annual rent (lnrent_olsen), and the local unemployment rate (FLBRA). Two are attempts to capture the impact of the distressed real estate inventory. One of these is a measure of the inventory of foreclosed properties awaiting sale back to the market via REO sales as a percent of the stock of owner-occupied housing with at least one mortgage (fore_cum_pct). Another is the volume of both regular and REO residential sales (lnallsales_CA). The inclusion of some other variables in the regression is driven by modest intuition and experimentation and step wise regression. These are the property tax rate for owners (proptaxrate), the home ownership rate (HOrate), and the size of the stock of owner-occupied housing with one or more mortgages (lnOOwM). Three measures of national variables are also included: the 1 year treasury rate (t1), the 10 year treasury rate (t10), and the national CPI (cpi_2009). Two methods are used to estimate the reduced form: simple regression with fixed effects for state and eight indicators of certain MSA group and a panel regression procedure in Stata (xtreg), which generates estimates of county level fixed effects. The key results are summarized in Table V-1.[7]

Consider, first, regression based estimates of the reduced form for the level of house prices (in log form) using the panel estimation procedure and county fixed effects. These results are contained in the first two columns of Table V-1; the first uses all counties for years 2005–2010; this includes 416 counties with data for 2005 and 439 with data for 2006 through 2010. The second column utilizes the balanced six year sample of 416 counties. The R^2 values and the RMSEs are virtually identical for each sample at 0.69 and 0.09, respectively. This is taken as a signal that little is lost by using the unbalanced sample for most of the specifications and the explanatory power of the equation is reasonable. The coefficient estimates in these two equations confirm the importance of some but not all of the traditional variables and highlight the role of others not typically included. The level of rents, the size of the stock of owner-occupied housing with at least one mortgage, the vacancy rate, and the unemployment rate perform similarly in both equations, are significant, and have the expected signs. Household income and the volume of sales show a much smaller than expected impact and the coefficient estimates are not statistically significant. The most dominant variable is a measure of the size of the distressed inventory (fore_cum_pct). Its coefficient estimate is about –5.5 in both equations and highly statistically significant, i.e. t statistics in excess of 15. Another strong performer is the property tax rate; higher tax rates, all else equal, are associated with lower house prices.

Now consider the estimates of the same two equations using ordinary regression (versus the panel method used in the first two columns), which are reported in columns 3 and 4. These estimates use dummy variables for states and MSAs to capture local market fixed effects. The results in column three use all counties for all years and column four focuses upon the balanced panel of counties for all years. The explanatory power of these two regressions is also high; R^2

[7] Some specifications were weighted by the counties 2000 population, but the main results seemed largely unchanged.

equals .88 and the RMSEs are .17 for the results both columns.[8] A comparison of the results for the two different estimation processes reveals a number of substantive differences.

First, and perhaps most importantly, many of the coefficient estimates prove to be quite sensitive to the estimation process and the treatment of fixed effects. Several of the traditional explanatory variables are now statistically significant that were not significant in the panel estimation process and some that were significant have different signs. Median household income has a strong, statistically significant, and positive impact on the level of house prices; the elasticity estimate is 1.1 and the t statistic is in excess of 37 for both specifications. The coefficient estimates using the panel process were small and insignificant. The coefficient estimates of sales volume and the home ownership rate are negative and significant using ordinary regression versus the finding of very small and insignificant estimates for these variables using the panel estimation process. Also, the level of annual rent adjusted by the Olsen rental price indexes is negatively related to the level of house prices; the elasticity estimates are statistically significant and negative and close to –.11 for both samples. The coefficient estimates of this variable were positive and significant in the panel estimation process. Similarly, the coefficient estimates of the vacancy rate are positive and significant using ordinary regression versus negative and significant in the panel estimation process. The estimates of the national information index (cpi2009) also changes signs and remains highly significant.

Second, there are also a number of strong similarities between the two sets of estimates. Of most importance is the performance of the distressed real estate inventory measure. This variable remains negative, substantial, and highly significant using the regression approach. The coefficient is, as expected, negative and highly significant. A one percentage point increase in this inventory reduces the level of house prices by 4 to 5 percent. The coefficient estimates for the unemployment rate and the property tax rate also remain highly significant and negative.

Third, both sets of estimates include three macro variables: the one year Treasury (t1); the 10 year Treasury (t10); and the national Consumer Price Index (cpi2009). These typically perform strongly in model with longer time series. The short-time series for this study and the tumultuous changes in monetary policy in the last several years reduces the likelihood of identifying a strong relationship with house prices. On the one hand, the coefficient estimate of the long term interest rate is highly significant and similar in magnitudes among the four sets of estimates. The shorter term interest rate goes from modestly negative and significant to substantially negative and significant. The estimates of the national consumer price index are statistically significant in both but opposite in signs.

These results suggest that any conclusions regarding the stability of reduced form model requires the choice of estimation technique. In terms of consistency of the estimates with a priori expectations, the regression results generally seem more plausible. This approach is thus used to do comparisons of coefficient estimates among various subgroups of the counties. The residuals from the regressions also seem to be most appropriate for the VAR equations that focus upon growth rate since these residuals are generated for each county and each year of the data. On the other hand, the residuals from the panel estimation process with the fixed effects may be a more

---

[8] A separate regression was estimated without any of the state and MSAgroup fixed effects using all counties. The R^2 for this equation was about .70 and the RMSE was about .26.

comprehensive and summary measure of the difficult to detect local market conditions over the 2005 through 2010 period. As such, the choice in this paper is to use the fixed effect estimates from the panel estimation as the dependent variable in the study of the fixed effects and the residuals from the regression models are used in the VAR analysis, both of which are discussed below.

Differences in the Reduced Form Estimates among Groups of Counties

Three different comparisons are considered. One consists of counties in the SAND states of Arizona, California, Florida, and Nevada, all of which were hit hard by the housing boom and bust, versus those in the nonSAND states. Another focuses upon differences between fast growing counties versus the slowest growing counties (Dpopgroup = 1 versus Dpopgroup =3).The third considers results based upon 2005 data versus those in 2010. Each is discussed in turn.

SAND versus nonSAND States. The SAND states are likely to show more signs of the impact of distress than the nonSAND states because they were hit so much harder by the current bubble-bust. For example, the 75 counties in the sample from the SAND states experienced declines in house prices of nearly 38 percent between 2005 and 2010 and 40 percent since the peak values in 2006 House prices peaked, on average, among the 344 counties in the nonSAND states in 2007. Prices in these states are 13 percent below their peak values and 7 percent below the 2005 values, on average. The percent of the stock of owner-occupied housing in Stage 3 of the foreclosure process (fore_cum_pct) was also about 2.4 percent in the SAND states and a full percentage higher than among counties in the nonSAND states, on average. These differences may lead to differences in the coefficient estimates of the reduced form not completely captured by the various fixed effect estimates. Columns 5 and 6 of Table V-1 contain the regression coefficient estimates for all counties and all years in the SAND and nonSAND states.

The explanatory power of the two equations for SAND and nonSAND is substantial; R^2 values are .88 and .93 for the SAND and nonSAND states, respectively. Many of the coefficient estimates for the SAND and nonSAND states are similar to one another and to the estimates for all counties (Column 4) such as income, the size of the stock of owner-occupied housing, the home ownership rate, the vacancy rate, and the unemployment rate. Several are distinctively different. Rent has a much larger negative and statistically significant coefficient than it does among the nonSAND states (–.43 versus –.03). Another striking difference is the coefficient of the fore_cum_pct, which is –8.82 among SAND states and –1.32 among nonSAND States. Clearly both the magnitude of the stock of distressed real estate and its impact upon the level of house prices are larger in the SAND states. Another difference is found for the volume of residential sales inclusive of both regular and REO sales; the coefficient of this measure of sales is negative and statistically significant in the SAND states and very small and insignificant among nonSAND states. Overall, these comparisons suggest a number of important differences in the quantitative relationship between the level of house prices and their core drivers between counties in these two groups of states.

Fast versus Slow Growing Counties. Another comparison focuses upon the fastest growing counties in terms of population (Dpopgroup = 3) versus those with the smallest rates of population growth (Dpopgroup = 1); see columns 7 and 8 of Table V-1. Glaeser and Gyourko (2010) and Follain (2011) show evidence that the supply elasticity of housing is smaller in slow growing areas relative to fast growing areas. Such a difference may affect the responsiveness of house prices to various shocks in exogenous variables. One example of this is the coefficient of income, which is nearly twice as large among the slow growing counties as among the fast growing (1.4 versus .78). The coefficient of rent is negative and significant for fast growing counties and near zero and insignificant among slow growing counties. The volume of sales has a negative coefficient for both groups, but the coefficient is statistically significant among the slow growing counties and insignificant and small among the fast growing counties. Again, the estimates of the coefficient of the foreclosure inventory differ substantially; the distressed real estate inventory has a substantially more negative impact in the fast growing counties than in the slow growing ones. The explanatory power of both equations is strong, especially in the slow growing counties ($R^2 = .95$).

2005 versus 2010. The last comparison in Table V-1 looks at the estimates of the reduced form equation using only data in 2005 to estimates using only data in 2010 for the same 416 counties in both years. Several striking differences appear. Income has a much smaller impact in 2010 than in 2005 (1.76 in 2005 versus 1.57 in 2010). The impact of the distressed real estate is statistically significant, substantial, and negative in both years, but its coefficient is much smaller in 2010 (–13.47 in 2005 versus –4.592 in 2010). Increases in rent have a substantial and negative impact upon the price level in both years, though its impact is larger in 2010 than in 2005 (–.368 in 2005 versus –.455 in 2010). The explanatory power of both equations is about the same, i.e. $R^2 = .75$.

The differences between the actual level of prices and the predicted values based upon the regressions (residuals) have the potential to provide insights about future price changes. According to the assumption embedded in the VEC approach, those places with relatively large residuals are more likely to see a slowdown in house price growth and vice versa. One test of their importance as bubble predictors is done by using the lagged values of these residuals in the VAR equation designed to predict changes in the level of house prices, which is discussed below. The separate regressions for 2005 and 2010 also offer an opportunity to examine how these residuals may have changed between these two years. The simple correlation between the two measures of the residuals is a very modest 38 percent; hence, they likely tell very different stories about which counties had overvalued properties in 2005 versus 2010. In fact, a graph showing both of these two series for all counties shows a wide range of differences between the residuals in 2005 and 2010. A more modest set of graphs provides insights about those at the top, bottom, and middle of the pack. Figure V-1 shows the counties with the ten smallest and the ten largest residuals in 2005 relative to the values of the residuals in 2010. The residuals are plotted on the vertical axis and the horizontal axis indicates the MSA in which the county is located. Figure V-2 shows the two sets of residuals for the 20 counties in the middle of the pack.

The comparison provides insights about the magnitude of the residuals and how they changed. Consider those with the largest residuals in 2005, which ranged from 34 to 53 percent of value and were identified as places where properties were undervalued relative to this definition of the fundamentals. These included six within the SAND states. The ten counties at the bottom showed overvaluations between 34 and 50 percent. None of these were in the SAND states and include places from various regions of the country. Both sets tended to show substantial changes between 2005 and 2010. The largest change among this group occurred for Broward County (Ft. Lauderdale MSA), which went from 34 percent to 8 percent. Eight of the ten had smaller residuals and were less overpriced in 2010 than in 2005. Two actually went the other direction: Orange County in the Durham, NC MSA and Kenosha County in the Duluth, MN MSA. Changes were substantial among those in the 20 in the middle of the pack as well. Thirteen saw increases in the residuals and seven decreases. Two of these are in Florida, but the rest are outside of the SAND states and are located in a wide variety of regions.

Where the Reduced Form Model Works Best. One indicator of how well the reduced form model fits local market experiences is obtained by examining the estimates of the fixed effects associated with each county. Perhaps the alpha-beta distinction used in the finance literature to describe the returns to a stock may be helpful here. The regression coefficients can be considered the "beta" of the model and indicate the responsiveness of local house prices to key drivers of house price. The fixed effect estimates generated by the panel estimation procedure are, in essence, the average value of the residuals from the regression for the six years of data conditional on the betas or estimated coefficients. In this sense, the fixed effect estimates generated by the panel estimation process can be considered the "alpha" term because it seeks to quantify the local market conditions specific to each county. The following discussion focuses upon one set of the fixed effects estimates to provide a sense of where the model coefficients (the betas) themselves capture the variation in house prices among markets and where the fixed effects (the alphas) seem to dominate. The discussion will focus on Figure V-3, which is a map of the fixed effects around the country and Table V-2, which focuses upon summary statistics about the ten counties where the fixed effects are very small and the beta component of the model was sufficient to explain the level of house prices in these places. The goal of this discussion is to provide a broad and descriptive overview of the fixed effects. A deeper dive that involves the estimation of a model to explain their variation is offered near the end of this section.

The average fixed effect generated by the panel regression is, as expected, about equal to zero (0.02) among the 416 counties with data in each of the six years. The fixed effects range from a low –.98 to a high of 1.33.[9] A positive value of the fixed effect for a county indicates that the explanatory variables in the regression model by themselves typically underestimated house prices for the county for the six years of data. The opposite is the case for counties with negative fixed effects. The magnitudes of these are quite substantial when placed in terms of logarithms since the dependent variable is the natural log of house values in a county. For example, the natural log of the average price of about $250,000 is 12.43 and an increase in the log from 12.43 by one (fixed effect = 1), generates a price to about $680,000. So a fixed effect of 1 or minus one is, in fact, quite substantial.

One way to capture the distribution of the fixed effect estimates is by a map (See Figure V-3). This shows the distribution of the fixed effect estimates for each of the counties in the sample; the darker red areas have the smallest fixed effects and the darker green areas have the largest ones. The coastal pattern seems most obvious: many of the coastal counties in California and the ones from Massachusetts down to Virginia have the highest fixed effects. The lowest ones are

[9] The estimates of this regression are in column 2 of Table V-1.

found in areas more distant from major coastal cities such as the eastern parts of California, Arizona, and parts of Florida.

Another perspective is offered by focusing on the counties with the smallest values of the fixed effects, which is where the estimated model coefficients do a good job of predicting average house values without the fixed effects or alphas. Table V-2 contains one dozen counties arranged in alphabetical order that meet these criteria. It contains selected traits of these twelve counties. The bottom row contains information about these traits for the entire sample.
Before discussing the traits of these twelve counties, note the last column on the right of Table V-2. This interquartile range statistic (IQR) is built upon the annual residuals from the ordinary regression model without the fixed effect option. Roughly speaking, the fixed effect estimate is an average of these annual residuals. The IQR statistic sheds light on whether the annual residuals are relatively tightly bunched or include some very large or small residuals in selected years that may have offsetting impacts and suggest a small fixed effect where, in fact, one does not exist. The largest IQR is 9 percent for Madison County in the St. Louis, MO MSA, but the others are largely less than or equal to 5 percent. This is taken as an indicator that most of these counties with relatively low fixed effects have consistently had relatively low annual residuals.

So where does the model fit very well? Three are west of the Mississippi River, but the other nine are east of it. The twelve counties are in eight different states; three are in Pennsylvania. The average property values in these counties are generally well below the average for the entire sample ($221,907); Baltimore County and St. Johns County are slightly above the average. Most of these twelve counties experienced much lower rates of appreciation in 2000–2005 than the entire sample (51 percent), but not all. Most experienced lower reductions in values relative to the entire sample, though St. Johns County and Washington County saw much larger declines than the average (–10.2 percent). Rent to value ratios and price to income ratios in these counties show wide variation around the sample mean values. Similar results are found regarding the foreclosure inventory, total population and the size of the housing stock. That is, these counties represent a relatively narrow range around the sample mean values, but they do not include the worst hit areas of the country such as those in the SAND states where the declines in house prices were the steepest in absolute terms and relative to what the model predicted.

Search for a Stable Model of Price Growth

The second key equation estimated in this paper is the VAR, which attempts to explain the growth in annual house price changes as measured by changes in the natural log of house prices (ln HP_CA). This equation estimates the relationship between annual price changes house prices as functions of the lagged value of the price change, the residual from the first stage reduced form equation, lagged changes in several key drivers of price change, and various fixed effects. The key drivers include: median household income; rent; sales; the foreclosure inventory; and the unemployment rate. Special attention is also given to three measures of preconditions prior to the bubble bust: high cost loans as a share of total mortgage lending in 2004 and 2005 (highcostshare); the growth in house prices between 2000 and 2005 (dfhfa2000_05); and the share of employment in construction in 2004 and 2005. As in the reduced form, an attempt is made to highlight the sensitivity of the model to the treatment of fixed effects. The results are contained in Table V-3. The first column only includes a fixed effect for the SAND states. The second column includes additional fixed effects for SAND states and the MSA groups; the first column omits these two sets of fixed effects. In both cases the three measures of conditions in 2004 and 2005 also serve as fixed effects and are included in both regressions. Attention is focused on the first column.

The coefficient of lagged house price growth (.25) suggests that positive momentum is present in the model, though this amount is much lower than those in longer time-series models where estimates are much closer to one. The coefficient of the lagged residual term (–.11) is significant and has the expected sign: larger residuals in the first stage VEC suggest that prices are above equilibrium and serve to slow down growth rates, all else equal. The coefficient of this term is near zero and insignificant when more fixed effects are included, which does not surprising. Again, these results offer strong confirmation of the role played by the foreclosure inventory in today's markets as the coefficient of the lagged value of this variable is substantially negative and highly significant (t = –12.73). Growth in the lagged volume of sales also has a statistically significant impact on house price growth but the sign is positive. Lagged income growth, surprisingly, has a negative coefficient, though the coefficient is statistically significant.

Of special interest in this table are the coefficients of the variables that capture certain key conditions prior to the bubble bust. Counties in which high cost mortgages were more prominent experienced slower growth than others. Those that experienced substantial house price growth between 2000 and 2005 also experienced slower growth. Both of these impacts are statistically significant. The share of employment in the construction industry, surprisingly, had a modest but positive impact on growth. The inclusion of additional fixed effects for SAND states and the msagroups generally reduces the size and the significance of the coefficient estimates, but the impacts of the sales, the foreclosure inventory and the indictors of prior conditions remain strong and significant. The explanatory power as noted by the R^2 of .37 is modest

In sum, the estimates of the VAR equation confirm the differential importance of three sets of factors on house price growth during the crisis.

- The first set represents the standard list of drivers in a VAR model of this type. Among the set of these factors considered, lagged changes in prices, the VEC residual, household income and annual rent are typically significant but their magnitudes are modest and sensitive to the treatment of fixed effects.
- The second set consists of one variable: the lagged value of the size of the distressed inventory in each county and each year of the data. The importance of incorporating this measure of the distressed inventory is strongly supported in all of the specifications and groupings of the data considered.
- The third set consists of three variables designed to represents preconditions in a county prior to the bust: the share of high cost lending in 2004–2005; the growth in house prices in 2000–2005; and the share of employment in 2004–2005. All three of these variables have substantial and statistically significant results. The share of high cost mortgages in 2004 and 2005 and the buildup in house prices between 2000 and 2005 are both estimated to have negative impacts upon house price growth in 2005 through 2010, which

as expected. The share of employment in construction prior to the bust is also substantial and statistically significant, but the sign is positive and different than expected.

Clearly, understanding what happened during this period of serious distress benefits from models that include all three sets of factors, which is a key goal of this paper and one necessary to gain a more complete picture of both stable and distressed housing markets.

Predictions for 2011

A critical trait of a good model is its ability to predict the future. Predictions are generated for 2011 using the estimated equation for price change in the second column of estimates in Table V-3, which included more location fixed effects than those in column 1. This can be done in the VAR context without any additional information about actual values of the variables in the model in 2011 because the VAR predictions for 2011 are generated by values of the explanatory variables in 2010. Furthermore, using data provided by Collateral Analytics allows for the calculation of the differences between the actual and predicted outcomes for 2011. Four exhibits are developed to highlight key aspects of the predictions relative to the actual outcomes.

The first exhibit is a map of the distribution of the model predictions errors for 2011 as defined by the percentage difference between the actual median sale price in 2011 and the predicted value (See Figure V-4). The counties with the darkest green stars are those in which the actual prices exceed predicted prices by the largest percentages; the opposite is the case for those with darker shades of red stars. One pattern seems clear: prices in many of the coastal areas of California, Florida, Texas, much of the South, and some of the coastal areas along the east coast exceed the levels predicted by the model for 2011. Actual prices tend to be below the model predictions in much of the rest of the country.

Table V-4 focuses upon counties in the two extremes of the distribution of the prediction errors and a group with the smallest differences between the actual and predicted values. The top third of the table includes 12 counties in which the observed level of prices exceeded the predicted values by 23 to 41 percent in 2011. These strongly outperformed the market, according to our VAR model. Only two are in SAND states, Placer County in the Sacramento area and St. Johns County in the Jacksonville, FL MSA. The others are spread along the East Coast, though two are in the Minneapolis area. The group of 12 counties in which the actual and predicted prices are quite close is described in the middle of the table. This middle group includes counties from all regions of the country. It includes only one from the SAND states, Santa Cruz County, which is a coastal area and may have been relatively unaffected by the boom and bust that hit many of the counties in California farther to the east.

The bottom third of Table V-4 contains those where the model grossly overestimates the level of prices in 2011. The potential of extreme outliers in modeling of this type is captured by St. Louis County where the model is off by more than 100 percent. Most of these are in the northeast and none are in the SAND states. An interesting case is Suffolk County, New York. This is a very expensive suburb within the larger New York metro area that has experienced substantial turmoil in its housing markets in recent years, though it did not seem to be as affected by the pre-bust growth in house prices or subprime lending as many within the SAND states. Yet, much turmoil exists. Follain, Miller, and Sklarz (2012a and 2012c) discuss some of the differences between Suffolk County, three others in New York, and four counties from California.[10] A key part of this study is the focus of attention upon the various ways to measure the distressed inventory. They highlight three stages: stage one consists of those with negative equity; stage two includes those in which serious delinquency has occurred; stage three, the one used in this study, includes those in which foreclosure has occurred and the property awaits sale in the REO market. This measure will be higher, all else equal, in those states in which foreclosure happens relatively quickly. Indeed, foreclosures take longer in New York State than in most states and much longer than in California. Hence, the observed measure of the distressed inventory in this model may understate the turmoil in the Suffolk area. In particular, other measures of the distressed inventory that incorporate the number of properties with negative equity may improve the predictive power of the model.

Two other looks at the distribution of the prediction errors are contained in Figure V-5 and Figure V-6. The first of these plots the actual versus the predicted values of house prices in 2011 and reveals a strong positive relationship between the two. No obvious or systematic errors associated with the level of prices are revealed. Figure V-6 looks at the distribution of the errors between the SAND states (right side) and the nonSAND states (left side). Among the SAND states, about two-thirds of the counties have actual house prices above the predicted values. The opposite is the case among the nonSAND states. All else equal, this may suggest that the recovery in the SAND states may be faster than what is captured by a model that focuses solely upon the years during the turmoil.

More importantly, a review of the prediction errors highlights the extraordinary variation among local market outcomes and the difficulty of capturing the many drivers of house prices in econometric models. Some potential areas for additional research suggested by these results are the inclusion of other measures of distressed real estate, more equations that capture the interaction between housing and other sectors and other dimensions of the housing market, e.g. sales.

Search for the Drivers of the Fixed Effects

Fixed effects play an important role in the model estimates discussed so far. The estimates themselves are sensitive to how the fixed effects are measured, the number of fixed effects included, and the estimation process used to estimate the fixed effects. They also have the potential to shed light on the local market conditions that affect them but are not integrated into the model. As noted above, the fixed effect estimates can be likened to the notion of "alpha" or firm specific drivers of the returns to a stock in the capital asset pricing model. In this sense, the following discussion seeks to shed light on specific local market conditions that may affect the sizes of the fixed effect estimations.

A variety of data have been collected for this purpose.

[10] Both are available at: http://www.proteckservices.com/homevalueforecast/hvf-lessons/

- Some are simple measures of distances between a county and a number of major metropolitan areas such as Los Angeles, Miami, New York, Chicago, and Boston, which are among the largest group of MSAs in terms of total employment. Providing a simple interpretation of the effects of these various distances is not easy, but they consistently appear as statistically significant drivers of the fixed effects.
- Other county specific and nontime varying physical measures have also been collected regarding the amount of water within a county and its total land area. From these, two density measures have been computed and they consistently add explanatory power to the simple regressions designed to explain the fixed effects. One is the ratio of the area occupied by water to the total land area of the county (water density) and the other is the total housing stock divided by the amount of land area in square miles (stock density). Another variable considered is the share of the MSA population located in the county (popshare).
- Another set includes a wide range of potentially relevant economic and demographic factors. These include the amount of employment in construction, the age of the population over 65, personal bankruptcies, and net migration.
- The last set is a unique set of variables related to the local regulatory policies and Saiz's measure of the supply elasticity of housing. These are developed and discussed in the work of Glaeser, Gyourko, and Saiz (2008) and Saiz (2010). The data include the Wharton Regulatory Index and specific supply elasticity estimates generated by Saiz.[11] RS Means Construction Cost Index data for 2008 are available in Glaeser, et al (2008) for selected MSAs. These have been assigned to counties based upon the MSA in which they are located.

Some experimentation with various combinations was conducted to generate a robust set of findings among the many possibilities, some of which is presented.

These variables are used to estimate a simple regression model to explain the variation in the fixed effect estimates among counties. The fixed effect estimates focused upon in this section are derived from the application of the xtreg or panel procedure to the full panel of counties for years 2005–2010 (column 2 in Table V-1). The results of the fixed effects regression are contained in Table V-5. The results of three different specifications are discussed. The first two columns use the 439 counties available in 2010 but do not include the last set of variables. Column 1 uses ordinary regression and column 2 uses stepwise regression. Column 3 also uses stepwise regression.

A common result in both columns 1 and 2 is that the share of the MSA's population in the county is associated with smaller fixed effects and those with higher water densities are associated with larger fixed effects. The main difference between these two columns is that three of the variables in column 1 do not enter the stepwise regression results in column 2: construction employment; the population over 65; and, the number of bankruptcies. Otherwise,

[11] These data are available at Saiz's web site: http://real.wharton.upenn.edu/~saiz/ .

the coefficient estimates and the statistical significance of the coefficients of the various distance measures and the measures of the physical aspects of the counties are quite similar.

The results in the last column shed light on the role of the various regulatory and elasticity information and their interaction with the other set of variables considered. The main results are these:

- The most important variable among this group appears to be the RS Means index of construction cost. The coefficient is highly significant (t=5.31). A ten thousand dollar increase in construction costs increases the fixed effect by 0.08;
- The Wharton regulatory index is positive and significant, but Saiz's elasticity estimate does not remain in the group suggested by stepwise regression, which likely reflects the fact that the Saiz elasticity variable is highly and negatively correlated with the Wharton index.
- The coefficient of the share of the MSA population in a county is statistically significant and negative;
- Water density remains positive and marginally significant;
- Unlike the results with all the counties, the coefficient of the share of the population above 65 is now significant and has a positive sign.
- The coefficient estimates for Net Migration and Bankruptcies remain about the same as in columns 1 and 2; and,
- Only two of the distance measures remain—the distances of a county to Los Angeles and to Miami.

The regression explains about 56 percent of the variation in the fixed effect estimates.
An interesting computation stems from the two distance variables that remain significant in the final column. With these estimates it is possible to estimate the joint effect of these distance measures on the fixed effect for a county. The following table shows some of these calculations for various states around the country.[12] The largest impacts are in the Northwest—Washington state and Northern California. The smallest impacts are in Florida.

State	Predicted Fixed Effect	Distance to LA	Distance to Miami
Washington	1.12	932	2451
California	0.94	257	2351
New York	0.92	2323	1146
Illinois	0.77	1606	1158
Florida	0.55	2235	223

These do not include the areas with the greatest combined distances to LA and Miami, which are those in Hawaii and Alaska. The largest combined distances are actually in Massachusetts and

[12] These equal .00021*distance to LA + .000377*distance to Miami.

New Hampshire, but weighted by the estimated coefficients these areas lead to smaller impacts on the fixed effects than in the Northwest.

These results are offered partly to shed some light on the drivers of the county fixed effects and the local market conditions that drive them. Obviously, the list of potential drivers is larger than the one considered. Other analysis not reported suggests that the estimates of the economic and demographic variables results are, not surprisingly, sensitive to the year of data used to define them. More data on one key driver of the fixed effects—the RS Means Index of construction costs—would allow expansion of this analysis to more counties. Despite these limitations, it is hoped that it raises awareness and provides some insights about the fixed effects and, more generally, the important role that local market conditions and policies have on movements in house prices.

A Discussion of Two Alternative Specifications of the Equilibrium Level of Prices

A critical part of bubble detection is the specification of the first stage VEC or reduced form equation to explain the level of prices and their deviation from equilibrium. The approach used above relies upon an equation with a variety of potential drivers as well as fixed effects. As discussed in the literature review above, two alternative and simpler approaches are also used in studies of this type: the Rent to Value ratio and the Price to Income ratio. This section examines evidence regarding the stability of these ratios during turbulent times.

Search for a Stable Rent to Value Ratio

In a competitive market, the rent to value ratio (R/V) is often expressed as follows:

> R/V = (risk free interest + expenses + property tax + risk premium – expected appreciation) Price of Housing.

Often this equation is used to infer an expected appreciation rate of the price of the housing stock if the other factors are assumed or held to be constant. Indeed, this was a major indicator used by Leamer (2003) and others to predict a potential house price bubble in the early 2000s. As noted by Follain and Giertz (2011) and in other work by Follain and Follain (2008), the validity of this approach to predict future house price appreciation rests on the strong assumption that other factors that influence R/V, such as the risk premium, are held constant.

A brief look at the distribution of this ratio among the 416 counties in the sample suggests wide variability in R/V among counties and years. The average for the entire sample is 4.5 percent, but it varies from about 0.6 percent to over 14.3 percent. It was lower in the early part of the sample, which many took to mean that either the risk premium was low or that expected appreciation was high. It was 20 percent higher by 2010 and the standard deviation of the ratio increased by over 20 percent.

Rent to Value for 416 Counties by Year				
Year	Mean	Std. Dev	Min	Max
2005	4.2%	1.9%	0.6%	12.2%
2006	4.1%	1.8%	0.6%	9.7%
2007	4.0%	1.8%	0.5%	10.5%
2008	4.4%	1.8%	0.5%	11.6%
2009	4.9%	2.1%	0.7%	13.6%
2010	5.2%	2.2%	0.6%	14.3%
All Years	4.5%	2.0%	0.5%	14.3%

A deeper exploration of the rent to value ratio is conducted by estimating a regression model to explain variations in the ratio among counties and years in the sample. Perhaps this will reveal a stable relationship between the R/V ratio and a core set of drivers. The results of three regressions to explain the R/V are presented in Table V-6. The first uses data for all years and counties. Column two uses 2005 data and column three 2010 data. The same set of core variables used in the reduced form equation above is used here as well. They explain about 77 percent of the variation in R/V using the pooled data and about 50 percent for each individual year.

Three stable results are obtained for the three specifications. First, median income has a statistically significant and negative relationship with the R/V ratio in all three regressions. Second, higher property tax rates are also found to increase the ratio, all else equal. The magnitude of the coefficient does vary among the three specifications quite a bit. The third consistent performer is one of particular interest—the measure of the foreclosure inventory. Indeed, the coefficient is positive and highly significant in all three regressions. This result is consistent with the notion that the distressed inventory may serve as a proxy for the risk premium; a larger distressed inventory increases the riskiness of an investment and, consequently, the rent to value ratio, all else equal. This interpretation is also consistent with a recent article by Follain, Miller, and Sklarz (2012b) in which they discuss variations in R/V among markets and offer data on R/V throughout the country. Their story and the one suggested by these regression results is that the variation in R/V is substantial and hard to interpret without much local information, which includes measures of the distressed inventory. The other coefficients vary among the three specifications in sign, size, and statistical significance.

Why does this matter? At a high level, the question is whether the R/V ratio is stable enough to be a good indicator of an investment opportunity, much as the P/E ratio is thought to be for stocks. But there is also an important policy debate underway—should the foreclosed properties be converted to rental properties. Our guess at this point is that the ratio can be a useful guide to such a policy, but better measures of the local conditions and the variability of this ratio within markets require much more attention.

Search for a Stable Price to Value Ratio

As noted earlier in the paper, many have also looked at the Price to Income (P/Y) ratio as a potential indicator of the equilibrium value of housing. Instances in which the ratio is relatively large suggest the potential of a bubble and vice versa. Though previous research and our own results demonstrate a strong relationship between various measures of income and the level of prices, the relationship is not always consistent enough to guarantee its reliability as a bubble detector. The following discussion reports on our attempts to estimate and identify a stable model of the price to income ratio.

As with the R/V ratio, a review of the simple statistics for the P/Y ratio reveals substantial variation over time and among counties. The ratio declines substantially between 2005 and 2010 from 4.7 to 3.7. The average for all years is about 4.3. Substantial variability is also suggested by the standard deviations and the extreme values.

Price to Income Ratios for 416 Counties by Year				
year	Mean	Std. Dev	Min	Max
2005	4.7	2.4	1.9	21.7
2006	4.8	2.4	1.9	21.1
2007	4.6	2.3	1.9	21.0
2008	4.0	1.8	2.0	19.0
2009	3.8	1.6	1.4	16.8
2010	3.7	1.7	1.5	19.6
All Years	4.3	2.1	1.4	21.7

The estimated equations are designed to see whether a stable relationship exists between the P/Y ratio and the same core set of economic variables used to examine the R/V ratio. One additional specification using all counties and year also includes a demographically based measure: the share of the housing stock owned by households 65 years of age and older (eldershareHVAL). The inclusion of this variable is motivated by the observation of Dreiman-Harter noted above, which highlighted the correlation between the average price to income ratio for the US and the percent of the population that is between 20 and 30 years of age. Estimates of the four P/Y equations are contained in Table V-7.

Four of the core set of variables are statistically significant in each column. Larger stocks of owner-occupied housing and higher vacancy rate increase the P/Y ratio. The larger the stock of owner occupied housing with a mortgage and the vacancy rate, the larger is the P/Y ratio. Larger volumes of residential sales and higher property tax rates reduce the P/Y ratio. The strongest consistent performer is, again, the measure of the size of the distressed real estate inventory. The size is dramatically larger in 2005 (column 2), which is likely driven by relatively small size of the inventory prior to the bust. Otherwise, the coefficient is statistically significant and substantial for the other three equations.

The fourth column includes one additional variable: eldershareHVAL, which is the share of the value of the entire owner-occupied housing stock in a county owned by households of age 65 or higher. This is computed using the ACS home owner estimates of value. As shown below, the mean value of the Elderly Share of House Value for the 416 counties over all years is 22 percent but varies between 6 and 54 percent among the counties (see the second row of the following table). For comparison purposes, the table below also reports in the first row the elderly share of the total population. This is typically much smaller than the elderly share of the housing stock, but they are highly correlated (82 percent).

Variable	Obs.	Mean	Std. Dv.	Min	Max
Elderly Share of Population	2496	12%	4%	5%	34%
Elderly Share of House Value	2496	22%	6%	6%	54%

The estimated coefficient of eldershareHVAL in Table V-7 (column 4) is 8.41 and is highly significant (t = 11.86). This is reported using regression for the entire sample without state fixed effects. Though this is not exactly the variable discussed by Dreiman-Harter, it seems to be opposite in sign to the trends suggested by data between the late 1970s and 2000. That is, here a larger presence of elderly home owners drives up P/Y; whereas the earlier trends suggested that higher P/Y would be positively associated with younger populations. More investigation of this issue may be considered.

In sum and as with the R/V ratio, some evidence is found to support a strong relationship between the ratio and a core set of variables for this data. However, the coefficient estimates are often sensitive to the specification and the years included. The role of distressed inventory is also confirmed. Nonetheless, it seems difficult to conclude that the ratio itself or a single equation does a very compelling job to focus attention on this ratio as a primary indicator of bubbles in the housing market.

Section VI: Conclusions, Possible Next Steps, and Some Policy Implications

One difference between this paper and from many recent studies of housing markets is its focus upon the period of the crisis itself. It reports on a search for the underlying structure of a model capable of predicting both the level of house prices and their rate of change during a period of great distress or turmoil in the housing market. Annual data at the county level from 2005 through 2010 are used for the study. This time period includes 2005 when house prices were approaching their peak values in this bubble-bust period. Peak values were reached in 2006 and 2007 in most places. House prices subsequently experienced substantial but widely varying declines from their peak values in most counties. Among the 416 counties focused upon in this study, prices in 2010 were, on average, 20 percent lower than their peak values. Prices were 41 percent below their peak values among counties in the SAND states, which is more than three times the size of the average declines observed in other states of 13 percent. The turmoil began after a large growth in house prices between 2000 and 2005 and has generated unprecedented volumes of mortgage foreclosures in the post WWII period and a substantial inventory of properties in some state of financial distress.

Focusing on this turbulent time period offers a number of challenges and opportunities. The opportunities include the availability of data not widely available for a longer time-series study at the county level and the chance to evaluate whether relationships estimated during relatively stable times endure during tumultuous times. This latter opportunity is likened to medical studies that incorporate the behavior of both healthy and unhealthy people in their data; insights about a particular remedy are likely to be more reliable when both types of people are studied. In this sense, focusing upon the period of distress in housing markets in 2005 through 2010 enriches the sample with relatively large numbers of unhealthy patients relative to much previous literature. Much previous literature about housing prices focuses on periods in which data over the past 30 or so years about housing markets are plentiful and, by the standard of the current crisis, dominated by relatively stable markets (healthy patients).

Three challenges to such a search are also highlighted in the introduction. One is the need to incorporate data to measure certain conditions prior to the period that may have influenced the degree of the turmoil to follow such as the buildup in prices in the 2000–2005 period and the surge in subprime mortgage lending. Another challenge is to expand the traditional list of house price drivers to include new measures that capture the secondary impacts of the housing bust such as the size of the distressed real estate inventory. These two challenges are in addition to the more obvious one—deciphering a stable and enduring structural model of house prices in the midst of tumultuous and relatively short time period highlighted by a wide variety of policy interventions and when much is still unknown about the nature of the new equilibrium and how close we are to it.

This final section offers an assessment of the search conducted in this paper for a stable structural model capable of predicting the level of house prices and being helpful in the detection of future house price bubbles. The last section provides a list of policy implications suggested by this work.

Results of the Search

1. The specification of the reduced form shows some resemblance to traditional models but some major adjustments are needed. The purpose of the reduced form equation in this and other papers that try to identify house price bubbles is to capture the spread or gap between the current levels of house prices in a market relative to their long-run equilibrium values. Estimates of the reduced form provided in this paper include some of the usual suspects as well as other variables not typically included in a reduced form (See Table V-1). The explanatory power of the equations used to estimate the reduced form equations is generally high; R^2 is .88 for the regression with state and MSAgroup fixed effects and about .70 when the panel estimation procedure with the county fixed effects is used. A number of the usual suspects perform reasonably well. For example, the elasticity of the level of prices with respect to median household income is about 1.1 and is highly significant using data for all counties; however, estimates of this elasticity vary between SAND states and nonSAND states (1.1 versus 0.99) and between fast growing and slow growing counties (.78 versus 1.4). Evidence of the effects of some other variables such as the level of rents, the size of the stock of owner-occupied housing with a mortgage, the overall housing vacancy rate, the volume of sales, the home ownership rate, and the

unemployment rate is also obtained, but the magnitudes and statistical significance of these effects typically vary widely among groups of counties and time periods. The most consistently strong performer in the reduced form model is not one of the usual suspects in models of house prices prior to 2005—the measure of the size of the distressed inventory. This is measured as the percent of the stock of owner occupied housing with a mortgage in which foreclosure has occurred and the property sits in an inventory of homes awaiting sale. A one percent increase in this inventory reduces prices by about 4 percent, on average among these counties and years; however, substantial variation exists in the size of this effect. The impact is about seven times larger in the SAND states versus the nonSAND states (8.82 percent versus 1.32 percent) and is also substantially larger among counties that were fast growing prior to the period of this study relative to the slower growing counties.

2. The size and distribution of the residuals from the reduced form equations for the level of house prices are substantial and vary widely among counties. The estimation of the reduced form devotes substantial attention to the measurement of county specific fixed effects. Here the evidence examined shows large and widely varying residuals among counties and years (See Figures V-1, V-2, and V-3). These residuals reveal that the actual level of prices exceeded the amount predicted by the variables in the regression by about 40 percent among a dozen or so counties in 2005 and they remain positive and substantial in 2010. The sizes of the residuals at the opposite end of the spectrum are also substantial and in many cases remain substantial in 2010. Many of the places with a good fit in 2005 are ones with widely varying residuals in 2010. A panel estimation method generates estimates of county fixed effects that also vary widely among counties. Overall, the results regarding the residuals and fixed effect estimates criterion suggest local market conditions not readily captured by the usual economic variables and even the size of the distressed inventory play a major role in driving the variation in the level of house prices among counties during this tumultuous period.

3. Similar conclusions emerge regarding the equations to predict house price growth. Indeed, the estimated model includes strong and statistically significant estimates of some of the usual suspects such as lagged price change, and the size of the residual from the reduced form, and the volume of residential sales, some others such as the changes in the rate of growth in household income and rent and changes in the unemployment rate do not seem to play important roles (See Table V-3). On the other hand, estimates of this model also strongly confirm the importance of incorporating measures of the distressed inventory and two preconditions often ascribed as causes of the current crisis—excessive growth rates in prices in 2000–2005 and substantial amounts of high cost mortgages in 2004–2005.

4. Out of sample predictions of price change for 2011 are highly correlated with actual outcomes, though substantial gaps between the actual and predicted values exist for a number of counties. The out-of-sample predictions of the level of house prices for 2011 prove to be highly correlated with the actual outcomes (See Figure V-5) though there are some noticeable patterns. The map in Figure V-4 indicates that actual prices in 2011 tend to above the model predictions in many coastal areas of the county and the opposite is true for much of the rest of the country. The distribution of these prediction errors also

differ between SAND and nonSAND states; the distribution of these errors is both more tightly concentrated and skewed toward more positive errors (see Figure V-6).

5. The ability to explain the fixed effects generated by the first stage VEC regressions suggests some patterns but much is left to understand. A set of county and MSA specific variables is used to explain the county fixed effect estimates generated by the panel estimation process. These fixed effect estimates are county specific and represent the average size of the reduced form model errors over the six years of the study. These fixed effects are important drivers of the explanatory power of the reduced form model and their inclusion can have substantial impacts on the estimated coefficients of the other variables in the model, especially the usual suspects. The evidence suggests a strong and positive connection between the fixed effects and two variables—construction costs and the Wharton Regulatory Index (See Table V-5)—among the 216 counties where information on both of these variables is available. A number of other variables also help explain the fixed estimates such as the water density of the county (positive) and a county's share of the MSA population (negative). The regression explains about 56 percent of the variation in these fixed effects. Further exploration of the local market conditions that drive these fixed effects seems to be fertile ground for additional research.
6. Rent to value and price to income ratios by themselves are too volatile to be reliable predictors of future house price bubbles. These two ratios are often used to help detect bubbles: lower rent to value ratios and higher price to income ratios relative to some historical average are thought to be indicative of a potential bubble. The evidence in this paper offers substantial evidence of widely varying ratios among counties and during the years of the study. Estimates of equations to predict these ratios do show some interesting and compelling findings regarding the usual suspects, but they also demonstrate the importance of inclusion of a measure of the distressed real estate inventory (Tables V-7 and V-7)

Overall, this search for a stable structure underlying changes in house prices can point to a number of stable and substantial findings. It also suggests areas in which future searches might look. One of these directions is to conduct even more geographically granular searches at, say, the zip code level or even at a more granular level. Alternative measures of the distressed inventory might also be considered. This paper focuses on the size of the foreclosure inventory, but recent work by Follain, Miller, and Sklarz (2012a, 2012c) discuss other stages of the distressed inventory that might be used such as the percent of the stock of housing with negative equity. Lastly, the results of this paper and many others also suggest where not to look for additional evidence—macro models focus upon trends in national indexes of the housing market. There has long been a tension between housing market analysts who use both local data and national data to discuss housing trends. However, the current crisis and the circumstances leading up to it should put to rest for some time any confidence in such models. Granted, there was a sense in which most of the country was impacted by a common shock associated with the financial crisis of 2008, but the evidence strongly supports the view that local market responses to this common shock are highly varied and dependent upon local market traits, some of which may be measurable and some of which are hard to incorporate into any model, especially macro type models.

Policy Implications

A lengthy and thorough discussion and development of the potential policy implications of the research reported in the current paper requires a separate or additional paper in order to describe the current state of the policy debate, the many types of policies being pursued, and some ways in which the results of the current paper are most relevant to the policy debate. Here a brief overview of some policy debates to which this paper may be relevant is offered.

The discussion begins by noting two high level discussions of possible policies by Ben Bernanke and other high level officials of the Fed. One is a speech by Bernanke (2010) in which he discusses some of the lessons learned about the discipline of economics from the current crisis. He offers his opinion that much of the problem dealt with what he labeled as economic management as opposed to two other categories of economics: economic science and economic engineering. He also offers suggestions about how the science of economics needs to improve. These include expanding our understanding of human behavior away from the standard model of the ever rational agent, paying more attention to extreme events, and building better bubble detection models. The results of this paper underscore the challenges of building bubble detection models and the need to focus more attention upon extreme outliers.

Dudley (2010) offers an interesting discussion of the many challenges associated with bubble detection, some of which is quite consistent with the themes of this paper and the earlier papers by Follain and Giertz. He discusses the need for better models of bubble detection and policies designed to slow a bubble before it bursts. He also says that "the policymaker needs to conduct a careful cost-benefit analysis, weighing how successful a particular policy might be in restraining the rise in asset prices versus how costly it would be to remain passive, letting the bubble grow and then potentially burst disruptively. Many factors will affect the outcome of this analysis including the magnitude of the potential asset bubble and whether the potential asset bubble is occurring in the equity or debt markets. In this analysis, the policymaker is likely to find that compared with equity market bubbles, credit market bubbles are more prone to generate higher costs when they burst. Thus, the benefits of preventing credit bubbles from forming and collapsing are likely to be higher." The results of the current paper regarding the role of distressed real state inventories are consistent with the notion the policymakers and others grossly underestimated the costs of this particular bubble since such variables were seldom, if ever, included in models of house price growth. Studying these costs and the evolution of the housing market will be helpful in future discussions about bubble detection and may shift the emphasis more towards avoidance. Detailed studies of local housing markets and the role of local market conditions can be helpful in this regard.

There are also a number of specific policy remedies being implemented that might be altered in light of the results of this study. These policies include Treasury's HAMP (mortgage modification program). Following up on the last comment, policies need to be more specific to local market conditions. For example, a HAMP program has the potential to promote improvement in the performance of the housing market and improve the lives of people deeply affected by foreclosure. However, it seems obvious that a single net present value model applied to all areas of the country and that relies upon models estimated prior to the bust and is not the best way to operate this program. Of course, incorporating the details of these local market conditions requires considerable research and more resources to build and monitor these programs, but the results of the current paper suggest that movement in this direction would be helpful.

The current environment is also marked by widely varying state and local policies being addressed to combat the crisis that need to be incorporated into models of house prices. These include local policies, the wealth of the affected population, local job prospects, and many others. For example, predicting the future of house prices in parts of New York State that have been hit very hard by the crisis need to take into account the evolving policies relating to Settlement Conferences, which have been changed twice during the crisis and altered again in 2011 due to budgetary considerations. Indeed, these policies have also revealed wide variability in their effectiveness among counties. Another example is California's ongoing debate about whether to allow dual tracking of both foreclosure and possible loan modifications. These policies likely matter and need to be incorporated into a comprehensive model in order to make credible long-run predictions of house prices.

Still another example of a potentially relevant policy in some areas includes Owner Education and Mortgage Counseling. Much work is being done to highlight the benefits of mortgage counseling at the time of the mortgage decision and among those facing the prospect of foreclosure. A rich but perhaps difficult to tap knowledge base also exists to shed light on the effectiveness of mortgage counseling: direct interviews with those seeking foreclosure counseling.[13] Another angle suggested by the research in this paper is to educate owners about the values of their properties. Is the gap between owner estimates of value and the "true" market value as large as is suggested by the data in this paper? The Fed is doing more work on this as well. Perhaps efforts to help provide owners with better information about the values of their homes may lead to more realistic buy and sell decisions.

Policies to develop better property valuation or appraisals for the purpose of property taxation are also relevant. One consistent finding in this paper is that the property tax rate proved to be an important driver of the level of prices and the R/Y and P/Y ratios. Incorporating the connection between property tax rates, the assessment of house values for property tax purposes, and the pace of the recovery is likely important and very challenging. Follain (2012) discusses some of the problems in New York State as part of a forum at the Rockefeller Institute of Government entitled "Local Fiscal Challenges and New York's Tax Cap." It reviews some of the large gaps between changes in the market values of properties and adjustments to the assessed values of the properties for the purpose of property taxes. Part of the problem is related to the reduction in the number of "regular" sales. Local assessors should be encouraged to move beyond some of their traditionally justified reluctance to use a broader array of sales and, in particular, REO sales. The

[13] For example, Follain has been involved in the local process and has had many discussions with those providing the counseling. An attempt was made to study New York State's Settlement Conference process. Funds were offered to incent some of those seeking counseling advice to participate in a survey. No one agreed to participate. This idea is on hold until we can find a way to tap into what people are doing, the kinds of assistance they are receiving, and the effectiveness of the advice.

larger appraisal industry is also struggling with how to deal with the problems of property valuation in declining markets.[14]

The paper closes with another policy implication of this and related work by Lo (2012) and Foote et al (2012): the need for economists in the current environment to be humble about our understanding of the causes of major bubbles, the extent of the fallout from this one, and policy remedies built upon structural models of the housing market. Though such models and econometric models of the type featured in this paper and the two previous ones by Follain and Giertz provide substantial insights about bubbles and busts, they do not generate precise estimates of the future or even a highly reliable guide to the potential for extreme and negative outcomes in years to come. This does not mean economists should stop trying to build such models, but rather that model building for this purpose should be done with the knowledge that the search for such models is complex and likely to continue for many years and until we reach the "new normal" and beyond.

[14] I was a subject matter expert on a panel developed to address this question for the Appraisal Foundation, who just released its report on the topic. A copy of the final report was posted in May and is now available to the public at: APB Valuation Advisory #3: *Residential Appraising in a Declining Market*.

References

1. Abraham, J. and Patrick Hendershott (1996). "Bubbles in Metropolitan Housing Markets." *Journal of Housing Research.*
2. Agarwal, Sumit, Brent W. Ambrose, Souphala Chomsisengphet, and Anthony B. Sanders (2012). "Thy Neighbor's Mortgage: Does Living in a Subprime Neighborhood Affect One's Probability of Default?" *Real Estate Economics*. http://www.areuea.org/members/publications/view_article.phtml?id=18378
3. Agarwal, Sumit, Brent W. Ambrose, and Yildiray Yildirim (2010). "The Subprime Virus: Theory and Evidence." http://ssrn.com/abstract=1612948
4. Arce, Óscar and David López-Salido (2011). "Housing Bubbles." *American Economic Journal: Macroeconomics.* http://www.aeaweb.org/articles.php?doi=10.1257/mac.3.1.212
5. Archer, Wayne R. and Brent C. Smith (2012). "The Influence of Exotic Mortgage Loans in the House Price Bubble: A Pilot Study of 20 Counties in Florida." *Presented at the 2012 AREUEA Meetings.*
6. Barlevy, Gadi and Jonas D. M. Fisher (2010). "Mortgage Choices and Housing Speculation." http://www.chicagofed.org/webpages/publications/working_papers/2010/wp_12.cfm
7. Bernanke, Ben (2010). "Implications of the Financial Crisis for Economics." *Federal Reserve Board of Governors.* http://www.federalreserve.gov/newsevents/speech/bernanke20100924a.htm
8. Blackley, Dixie M. and James R. Follain (1991). "An Econometric Model of the Metropolitan Housing Market." *Journal of Housing Economic*s.
9. Campbell, Sean D., Morris A. Davis, Joshua Gallin, and Robert F. Martin (2009), "What moves housing markets: A variance decomposition of the rent–price ratio." *Journal of Urban Economics, Volume 66,* Issue 2, pp. 90–102.
10. Campbell, John Y., Stefano Giglio, Parag Pathak (2009). "Forced Sales and House Prices." http://www.nber.org/papers/w14866
11. Carrillo, Paul E., Dirk W. Early, and Edgar O. Olsen (2010). "A Panel of Price Indices for Housing, Other Goods, and All Goods for All Areas in the United States 1982–2008." http://www.virginia.edu/economics/Workshops/olsen/CEOFinalDecember2010.pdf
12. Case, Karl E. and Robert J. Shiller (2004). "Is There a Bubble in the Housing Market?" *Brooking Papers on Economic Activity.* http://cowles.econ.yale.edu/P/cp/p10b/p1089.pdf
13. Chen, Celia*,* Andres CarBaCho-Burgos, SunaYana Mehra, and Mike Zoller (2011). "Moody's Analytics Case-Shiller Home Price Index Forecast Methodology." *Moody's Analytics.* http://www.economy.com/home/products/samples/case-shiller-methodology.pdf
14. Clayton, Jim, Norman Miller & Liang Peng (2010). "Price-volume Correlation in the Housing Market: Causality and Co-movements," *Journal of Real Estate Finance and Economics.*

15. Coleman, Major D., Michael LaCour-Little, and Kerry D. Vandell (2008). "Subprime Lending and the Housing Bubble: Tail Wags Dog?" *Journal of Housing Economics.* http://ssrn.com/abstract=1262365

16. Dokko, J., Doyle, B. M., Kiley, M. T., Kim, J., Sherlund, S., Sim, J. and Van Den Heuvel, S. (2011). "Monetary Policy and the Global Housing Bubble," *Economic Policy*.

17. Dreiman-Harter, Michelle (2000). "Using the Price to Income Ratio to Determine the Presence of Housing Price Bubbles." http://www.fhfa.gov/webfiles/1071/Focus4Q00.pdf

18. Dreiman-Harter, Michelle (2004). "Drawing Inferences about Housing Supply Elasticity from House Price Responses to Income Shocks," *Journal of Urban Economics.* http://www.fhfa.gov/webfiles/1262/workingpaper032.pdf

19. Dudley, William C. (2010). "Asset Bubbles and the Implications for Central Bank Policy." *Speech transcript*. http://www.newyorkfed.org/newsevents/speeches/2010/dud100407.html

20. Engel, Kathleen C. and Patricia A. McCoy (2011). The Subprime Virus: Reckless Credit, Regulatory Failure, and Next Steps, *Oxford University Press.*

21. Follain, James R. (2009). "A Look at Seller Behavior and Outcomes during Uncertain Times." *Collateral Analytics.* https://collateralanalytics.com/15/

22. Follain, James R. (2011). "A Study of Real Estate Markets in Declining Cities." *Research Institute for Housing America, Special Report.* http://www.housingamerica.org/Publications/AStudyofRealEstateMarketsinDecliningCities.htm

23. Follain, James R. (2012). "The Impact of the Great Recession on the Property Tax in New York State." http://www.rockinst.org/pdf/government_finance/2012-03-01-Recession_Property_Tax.pdf http://www.rockinst.org/pdf/government_finance/2012-03-01-Recession_Property_Tax.pdf

24. Follain, James R. (forthcoming). "The Search for Capital Adequacy in the Mortgage Market: A Case of Black Swan Blindness." *International Journal of Housing Markets and Analysis.*

25. Follain, James R. and Barbara Follain (2007). "Are We There Yet?" http://www.ficonsulting.com/node/17

26. Follain, James R. and Barbara Follain (2008). "Under the Hood of New Fed Plan for Homeowners." http://www.ficonsulting.com/node/17

27. Follain, James R. and Seth H. Giertz (2011a). "Using Monte Carlo Simulations to Establish a New House Price Stress Test." *Journal of Housing Economics.*

28. Follain, James R. and Seth H. Giertz (2011b). "A Look at US House Price Bubbles from 1980–2010 and the Role of Local Market Conditions." *Lincoln Institute of Land Policy working paper.*

29. Follain, James R. and Seth H. Giertz (2012). "Predicting House Price Bubbles and Busts with Econometric Models: What We've Learned. What We Still Don't Know." *Lincoln Institute*

of Land Policy Working Paper. *http://www.lincolninst.edu/pubs/1928_A-Look-at-US-House-Price-Bubbles-from-1980-2010-and-the-Role-of-Local-Market-Conditions*

30. Follain, James R. Jr., and Stephen Malpezzi (1981). “Are Occupants Accurate Appraisers?” *Review of Public Data Use*.

31. Follain, James R., Norman Miller, and Michael Sklarz (2012a). “Measuring the Size of the Inventory of Distressed Real Estate: The Drain Is Still Clogged.” *Lessons from the Data.* http://www.proteckservices.com/hvf-lessons-from-the-data/measuring-the-size-of-the-inventory-of-distressed-real-estate-the-drain-is-still-clogged/

32. Follain, James R., Norman Miller, and Michael Sklarz (2012b). “Rent to Value Ratio: The Economics of Rental Property.” *Lessons from the Data.* http://www.proteckservices.com/hvf-lessons-from-the-data/rent-to-value-ratio-the-economics-of-rental-property/

33. Follain, James R., Norman Miller, and Michael Sklarz (2012c), “Negative Equity: Stage One of Distressed Real Estate Inventory.” *Lessons from the Data.* http://www.proteckservices.com/hvf-lessons-from-the-*data/negative-equity*-stage-one-of-distressed-real-estate-inventory/

34. Follain, James R. and Mike Sklarz (2005). “Pricing Market Specific Bubbles,” *Mortgage Banking Magazine.*

35. Follain, James R. and Orawin Velz (1995). “Incorporating the Number of Existing Home Sales into a Structural Model of the Market for Owner-Occupied Housing.” *Journal of Housing Economics.*

36. Foote, Christopher L., Kristopher S. Gerardi, and Paul S. Willen (2012). “Why Did So Many People make So Many Ex Post Bad Decisions? The Causes of the Foreclosure Crisis” *Federal Reserve Bank of Boston, Public Policy Discussion Paper.* http://www.bostonfed.org/economic/ppdp/2012/ppdp1202.html

37. Glaeser, Edward. L. and Joseph Gyourko (2005). “Urban Decline and Durable Housing,” *Journal of Political Economy.*

38. Glaeser, E., Joseph Gyourko and Raven E. Saks (2005). “Why Have Housing Prices Gone Up?” *American Economic Review.*

39. Glaeser, Edward L., Joseph Gyourko, and Albert Saiz (2008). “Housing supply and housing bubbles.” *Journal of Urban Economics.* http://ideas.repec.org/a/eee/juecon/v64y2008i2p198-217.html

40. Glaeser, Edward L., Joshua D. Gottlieb, and Joseph Gyourko (2010). “Can Cheap Credit Explain the Housing Boom?” *NBER Working Paper*. http://www.nber.org/papers/w16230

41. Goetzmann, William N., Liang Peng, and Jacqueline Yen (2009). “The Subprime Crisis and House Price Appreciation.” *Yale ICF Working Paper.* http://ssrn.com/abstract=1340577

42. Goodman, Allen C. and Thomas G. Thibodeau (2008). “Where are the speculative bubbles in US housing markets?” *Journal of Housing Economics.*

43. Goodman, John L. and John B. Ittner (1992). "The Accuracy of Home Owner's Estimates of House Value." *Journal of Housing Economics.*

44. Gyourko, Joseph, Albert Saiz, and Anita Summers (2008). "A New Measure of the Local Regulatory Environment for Housing Markets: The Wharton Residential Land Use Regulatory Index." *Urban Studies.* http://real.wharton.upenn.edu/~gyourko/WRLURI/The%20Wharton%20Zoning%20Regulation%20Index-July%202,%202007.pdf

45. Hendershott, Patric, Robert Hendershott, and James Shilling (2010). "The Mortgage Finance Bubble: Causes and Corrections." *Journal of Housing Research.* http://economics.kenyon.edu/melick/Econ391/A-Mortgage/HendershottMortgageFinanceBubble.pdf

46. Kiefer, Hua and Leonard C. Kiefer (2009). "The Co-Movement of Mortgage Foreclosure Rate and House Price Depreciation: A Spatial Simultaneous Equation System." http://www.areuea.org/conferences/papers/download.phtml?id=2570

47. Kiel, Katherine A. and Jeffery E. Zabel (1999). "The Accuracy of Owner-Provided House Values: The 1987-1991 American Housing Survey." *Real Estate Economics.* http://www.questia.com/read/1G1-55282179/the-accuracy-of-owner-provided-house-values-the-1978-1991

48. Kim, Sei-Wan and Radha Bhattacharya (2009). "Regional Housing Prices in the USA: An Empirical Investigation of Nonlinearity," *Journal of Real Estate Finance and Economics.* http://www.springerlink.com/content/q48r554917020764/fulltext.pdf

49. Kishor, N. Kundan and James Morley (2010). "What Moves the Price–Rent Ratio: A Latent Variable Approach"? *Working paper.* http://apps.olin.wustl.edu/MEGConference/Files/pdf/2010/48.pdf

50. LaCour-Little, Michael and Jing Yang (2010). "Pay Me Now or Pay Me Later: Alternative Mortgage Products and the Mortgage Crisis." *Real Estate Economics.*

51. LaCour-Little, Michael, Eric Rosenblatt, and Vincent Yao (2010). "Home Equity Extraction by Homeowners: 2000–2006." *Journal of Real Estate Research.*

52. Lai, Rose Neng and Robert Van Order (2010). "Momentum and House Price Growth in the U.S. Anatomy of a Bubble." *Real Estate Economics.*

53. Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2012). "The Housing Market(s) of San Diego." *NBER Working Paper.* http://www.nber.org/papers/w17723

54. Leamer, Edward (2003). "Bubble Trouble? Your Home Has a P/E Ratio Too." *UCLA Anderson Forecast Report*. http://www.uclaforecast.com/ and the update: http://www.anderson.ucla.edu/documents/areas/ctr/forecast/PE_ratio_update.pdf

55. Lo, Andrew W. (2012). "Reading about the Financial Crisis: A Twenty-One-Book Review." *Journal of Economic Literature.* http://www.aeaweb.org/articles.php?doi=10.1257/jel.50.1.151 or http://ssrn.com/abstract=1949908

56. Lu, Lexian and Wei Wang (2010). "Similar Cities vs. Nearby Cities: A Spatial Econometric Approach to MSA Level Housing Returns." *Working paper.* https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=FEMES09&paper_id=647

57. Malpezzi, Stephen (1999). "Simple Error Correction Model of House Prices," *Journal of Housing Economics.*

58. Mian, Atif, Amir Sufi, and Francesco Trebbic (2011). "Foreclosures, House Prices, and the Real Economy" *NBER Working Paper.* http://www.nber.org/papers/w16685.pdf

59. Miller, Norman G., Liang Peng, and Michael Sklarz (Forthcoming)." House Prices and Economic Growth." *Journal of Real Estate Finance and Economics.* http://ssrn.com/abstract=1427570

60. Miller, Norman and Liang Peng and Michael Sklarz (2011) "The Economic Impact of Anticipated House Price Changes—Evidence from Home Sales." *Real Estate Economics.*

61. Morgan, Donald P., Benjamin Iverson, and Matthew Botsch (Forthcoming). "Subprime Foreclosures and the 2005 Bankruptcy Reform" *Economic Policy Review*, Federal Reserve Bank of New York.

62. Saiz, Albert (2010). "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics*. http://qje.oxfordjournals.org/content/125/3/1253.abstract

63. Taleb, Nicolas (2007). The Black Swan: The Impact of the Highly Improbable, *Random House.*

64. Wheaton, William C. and Gleb Nechayev (2008). "The 1998–2005 Housing 'Bubble' and the Current 'Correction': What's Different this Time?" *Journal of Real Estate Research.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1028895

65. Zhou, Jian (2010). "Testing for Cointegration between House Prices and Economic Fundamentals," *Real Estate Economics.*

Table III-1: Summary of House Price Variables						
year		HP_ACS	HP_CA	HP_CA2	MSP	HP_CA
						Balanced
2005	N	416	416	329	416	416
	Mean	$ 218,659	$ 244,654	$ 207,982	$ 231,111	$ 244,654
2006	N	439	439	402	439	416
	Mean	$ 235,035	$ 252,936	$ 222,571	$ 236,869	$ 258,949
2007	N	439	439	423	439	416
	Mean	$ 242,116	$ 252,337	$ 235,649	$ 239,170	$ 258,411
2008	N	439	439	426	439	416
	Mean	$ 239,991	$ 229,423	$ 234,254	$ 213,843	$ 234,797
2009	N	439	439	426	439	416
	Mean	$ 224,198	$ 206,313	$ 214,607	$ 184,825	$ 211,098
2010	N	439	439	431	439	416
	Mean	$ 217,508	$ 202,530	$ 193,424	$ 187,666	$ 207,370
Total	N	2611	2611	2437	2611	2496
	Mean	$ 229,681	$ 231,248	$ 218,367	$ 215,444	$ 235,880

Table III-2: Cumulative Percentage Changes in the Indexes					
Year	HP_ACS	HP_CA	HP_CA2	MSP	FHFA
	Cumulative Percentage Change since 2005				
2006	10%	7%	14%	5%	9%
2007	14%	7%	24%	6%	11%
2008	16%	0%	24%	-3%	7%
2009	10%	-8%	14%	-13%	2%
2010	7%	-11%	3%	-13%	-3%

Table III-3: House Prices among Groups of Counties (2005 through 2010)						
	HP_CA					
Year	All Counties	Non SAND	SAND	Slow Growing Dpop=1	Medium Growing Dpop=2	Fast Growing Dpop=3
2005	$ 244,654	$ 218,193	$ 364,964	$ 279,679	$ 240,623	$ 219,960
2006	**$ 252,936**	$ 230,471	**$ 388,428**	$ 291,587	**$ 252,045**	**$ 238,604**
2007	$ 252,337	**$ 233,828**	$ 370,184	**$ 295,120**	$ 251,009	$ 235,222
2008	$ 229,423	$ 222,700	$ 289,798	$ 277,543	$ 223,904	$ 209,719
2009	$ 206,313	$ 206,484	$ 232,075	$ 255,075	$ 201,839	$ 183,647
2010	$ 202,530	$ 202,953	$ 227,457	$ 252,332	$ 197,784	$ 179,407
Pct Change since '05	-17%	-7%	-38%	-10%	-18%	-18%
Pct Change since peak	-20%	-13%	-41%	-14%	-22%	-25%
Peak	$ 252,936	$ 233,828	$ 388,428	$ 295,120	$ 252,045	$ 238,604
	$ 202,530	$ 202,953	$ 227,457	$ 252,332	$ 197,784	$ 179,407
HP_CA is the CA Measure of Collateral Analtyics's Median Sales Price per square foot times 1,750 square feet.						

Table III-4: HVAL Gap			
year		nonSAND States	SAND States
2005	HVALgap	-9%	-13%
	fore_cum_pct	0%	0%
	dmsft_pct_adj	0%	0%
2006	HVALgap	-7%	-7%
	fore_cum_pct	1%	0%
	dmsft_pct_adj	6%	9%
2007	HVALgap	-4%	-2%
	fore_cum_pct	1%	1%
	dmsft_pct_adj	8%	3%
2008	HVALgap	3%	15%
	fore_cum_pct	1%	2%
	dmsft_pct_adj	4%	-18%
2009	HVALgap	7%	24%
	fore_cum_pct	1%	2%
	dmsft_pct_adj	-2%	-35%
2010	HVALgap	8%	18%
	fore_cum_pct	2%	2%
	dmsft_pct_adj	-4%	-39%
HVAL_gap = ACS Homeowner estimate of value less the HP_CA market based value			
fore_cum_pct = Foreclosure inventory relative to the stock of OO Housing w/mortgage			
dmsft_pct_adj= Percent change in the Median Sale Price per square foot.			

Table III-5: Indicators of Rent, Rent to Value Ratios, and Price to Income Ratios					
2005	Number of Counties	Mean	Std. Dev	Min	Max
ACS Rent (annual)	390	$ 8,999	$ 2,022	$ 5,412	$ 16,728
ACS Rent_Olsen Adjusted	390	$ 8,158	$ 1,085	$ 5,289	$ 12,260
FMR 2 Bedroom	390	$ 9,145	$ 2,396	$ 5,064	$ 18,468
FMR 3 Bedroom	390	$ 12,101	$ 3,206	$ 6,084	$ 24,660
Rent to Price Ratio (Olsen-Adj Rent to HP_CA)	390	4.3%	1.9%	0.6%	12.2%
Rent to Price Ratio (ACS Rent to HP_ACS)	390	4.1	2.0	2.0	12.8
Price to Income Ratio (HP_CA to ACS Median Income)	390	4.6	2.4	1.9	21.7
2010					
ACS Rent (annual)	390	$ 10,506	$ 2,377	$ 6,672	$ 19,176
ACS Rent_Olsen Adjusted	390	$ 8,547	$ 1,113	$ 5,702	$ 12,530
FMR 2 Bedroom	390	$ 10,866	$ 2,932	$ 6,648	$ 21,120
FMR 3 Bedroom	390	$ 14,329	$ 3,885	$ 7,980	$ 29,676
Rent to Price Ratio (Olsen-Adj Rent to HP_CA)	390	5.2%	2.1%	0.6%	14.3%
Rent to Price Ratio (ACS Rent to HP_ACS)	390	3.9	1.5	2.1	13.4
Price to Income Ratio (HP_CA to ACS Median Income)	390	3.7	1.7	1.5	19.6

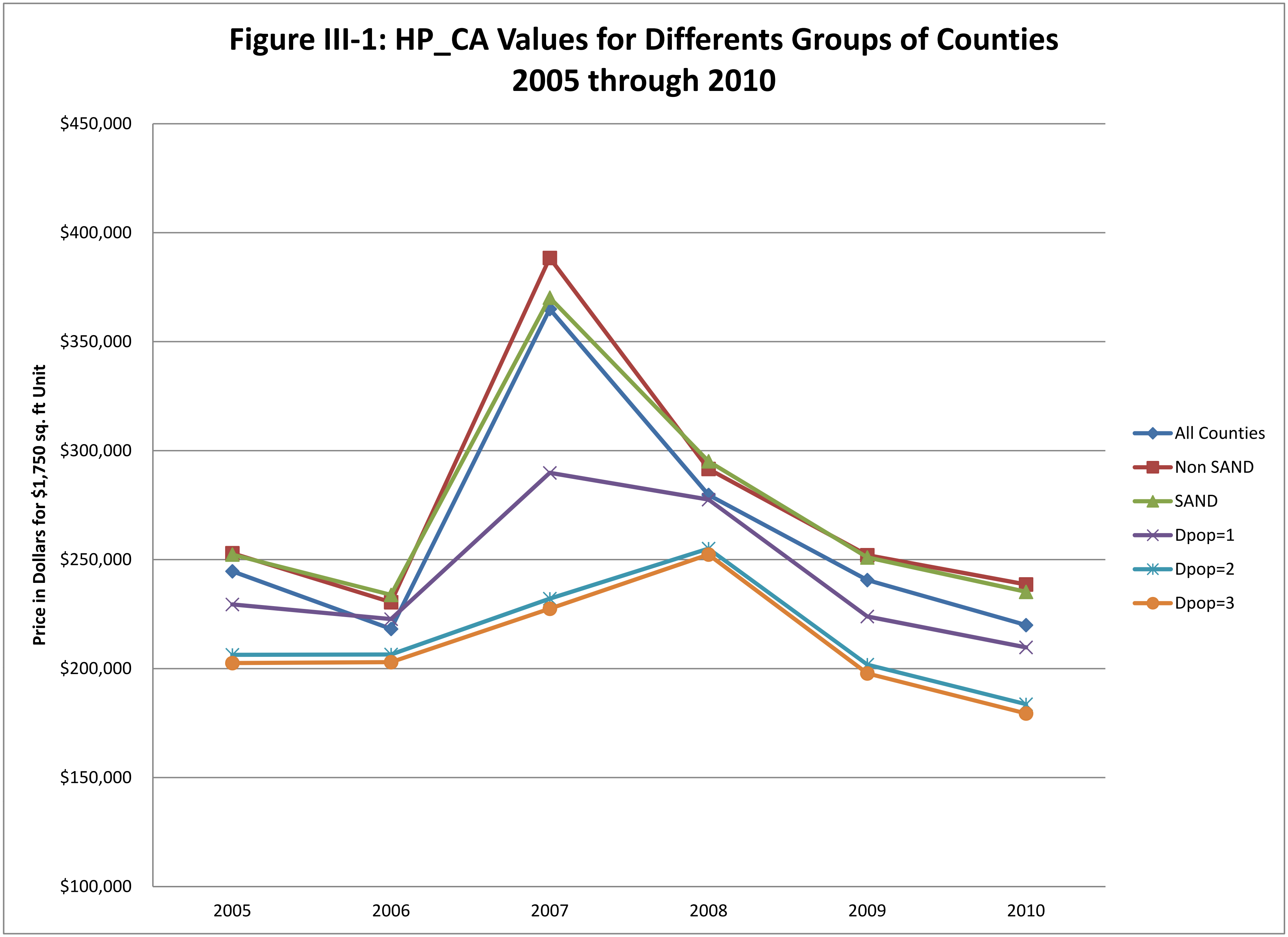


Figure III-1: HP_CA Values for Differents Groups of Counties
2005 through 2010
Price in Dollars for $1,750 sq. ft Unit
$450,000
$400,000
$350,000
$300,000
$250,000
$200,000
$150,000
$100,000
2005
2006
2007
2008
2009
2010
All Counties
Non SAND
SAND
Dpop=1
Dpop=2
Dpop=3

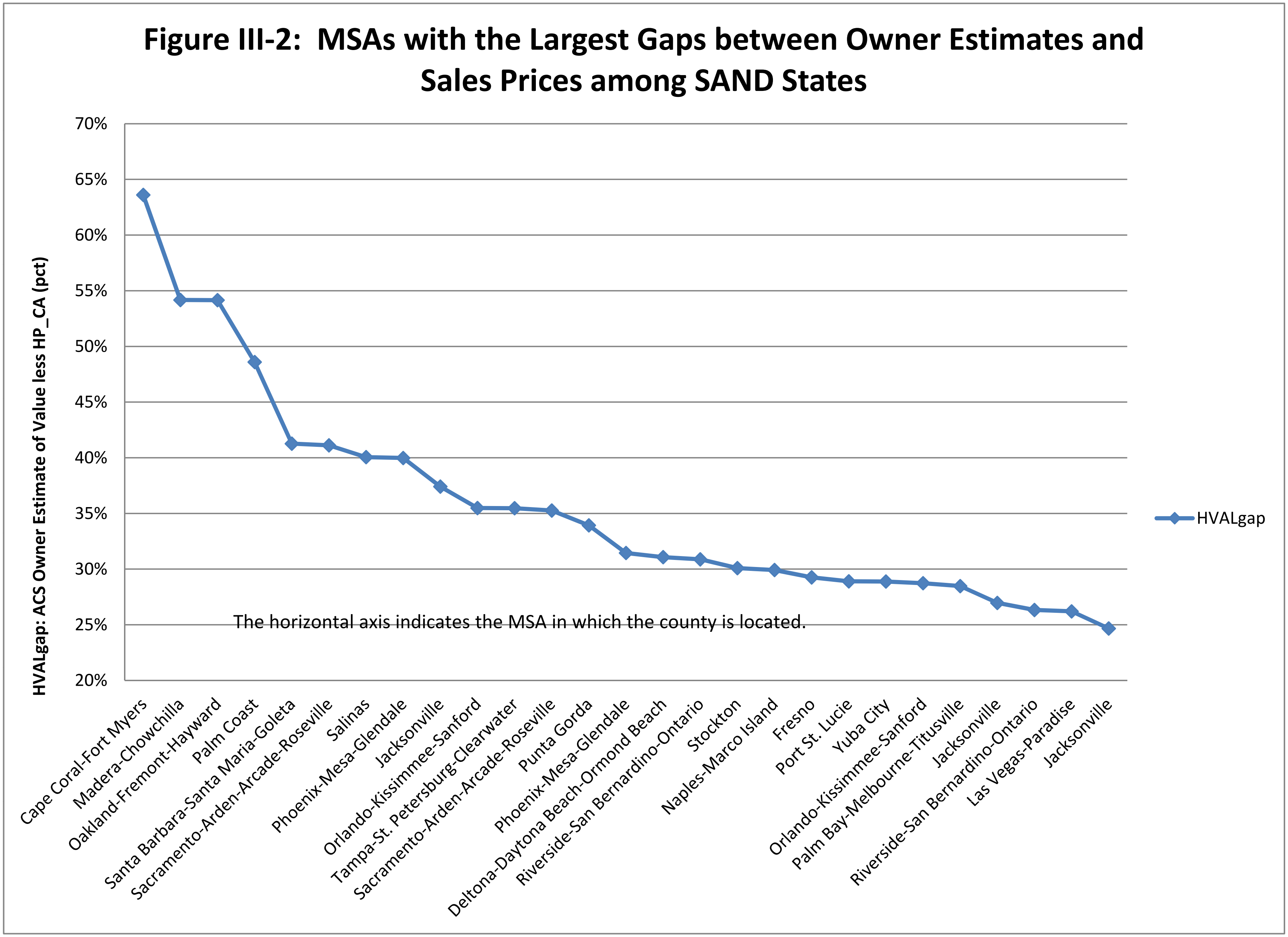


Figure III-2: MSAs with the Largest Gaps between Owner Estimates and Sales Prices among SAND States
HVALgap: ACS Owner Estimate of Value less HP_CA (pct)
70%
65%
60%
55%
50%
45%
40%
35%
30%
25%
20%
The horizontal axis indicates the MSA in which the county is located.
HVALgap
Cape Coral-Fort Myers
Madera-Chowchilla
Oakland-Fremont-Hayward
Palm Coast
Santa Barbara-Santa Maria-Goleta
Sacramento-Arden-Arcade-Roseville
Salinas
Phoenix-Mesa-Glendale
Jacksonville
Orlando-Kissimmee-Sanford
Tampa-St. Petersburg-Clearwater
Sacramento-Arden-Arcade-Roseville
Punta Gorda
Phoenix-Mesa-Glendale
Deltona-Daytona Beach-Ormond Beach
Riverside-San Bernardino-Ontario
Stockton
Naples-Marco Island
Fresno
Port St. Lucie
Yuba City
Orlando-Kissimmee-Sanford
Palm Bay-Melbourne-Titusville
Jacksonville
Riverside-San Bernardino-Ontario
Las Vegas-Paradise
Jacksonville

	Table V-1: Estimates of the Reduced Form for House Prices (VEC)									
Method	XTREG with fixed effects	XTREG with fixed effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects	Ordinary Regression with State and MSA Fixed Effects
Selected Counties	All	Balanced	All	Balanced	All SAND	All NonSAND	Dpop==1	Dpop==3	All in 2005	All in 2010
Years	2005-2010	2005-2010	2005-2011	2005-2010	2005-2014	2005-2015	2005-2016	2005-2017	2005	2010
	Dependent Variable: Natural Log of the Level of House Prices as measured by lnHVAL_CA									
Exogenous Variables	**1**	**2**	**3**	**4**	5	6	**7**	**8**	**9**	**10**
Median Household Income	0.01	0.00	1.11	1.10	0.99	1.11	1.40	0.78	1.76	1.57
	0.28	-0.02	37.82	37.54	14.36	35.21	27.89	16.67	20.65	23.16
Median Rent	0.16	0.18	-0.11	-0.11	-0.42	-0.03	-0.21	0.03	-0.37	-0.46
	3.30	3.64	-3.02	-2.80	-5.29	-0.77	-3.53	0.51	-3.37	-4.83
No. of OO HU (OOwM)	0.18	0.21	0.08	0.08	0.16	0.07	0.12	0.04	0.09	0.08
	3.48	3.74	11.47	11.10	6.85	9.84	10.10	4.04	3.66	3.07
Housing Vacancy Rate	-0.57	-0.68	0.76	0.89	1.00	0.70	0.63	0.80	1.66	0.74
	-3.97	-4.59	8.73	10.07	6.62	7.35	3.28	7.28	5.19	3.44
Regular and REO sales	0.00	0.00	-0.02	-0.02	-0.15	-0.01	-0.06	-0.01	0.03	-0.02
	0.30	0.17	-4.33	-4.19	-7.41	-1.84	-5.54	-0.88	1.71	-1.14
Foreclosure Inventory (% of OOwM)	-5.46	-5.59	-4.10	-4.86	-8.82	-1.32	-1.20	-7.74	-13.47	-4.59
	-15.47	-15.71	-10.90	-12.53	-10.64	-3.43	-1.90	-11.64	-5.03	-5.76
Homeownership rate	-0.04	0.01	-1.29	-1.31	-0.97	-1.26	-1.48	-1.22	-2.31	-2.00
	-0.29	0.07	-24.39	-24.48	-7.93	-23.09	-19.16	-13.14	-13.96	-15.07
Property Tax Rate	-57.09	-56.49	-38.29	-37.25	-25.86	-39.17	-29.87	-47.63	-29.38	-20.69
	-24.45	-24.11	-25.91	-24.72	-3.82	-27.99	-16.47	-13.11	-10.83	-8.87
Unemployment Rate	-0.02	-0.02	-0.02	-0.02	-0.01	-0.01	0.01	-0.02	0.02	0.00
	-10.98	-10.83	-8.50	-8.25	-3.27	-2.36	1.18	-5.27	1.84	-0.43
One year treasury rate	-0.003	-0.004	-0.02	-0.01	-0.01	0.00	0.00	-0.01		
	-1.25	-1.56	-3.30	-3.18	-0.62	-0.70	-0.55	-0.86		
Ten year treasury rate	0.03	0.03	0.05	0.05	0.10	0.03	0.05	0.05		
	5.42	5.60	4.82	4.76	4.56	3.09	3.50	3.48		
CPI Index base 2009	0.76	0.81	-0.93	-0.84	-2.78	-0.84	-1.52	-0.37		
	8.13	8.49	-7.21	-6.53	-7.67	-6.61	-7.93	-1.87		
Constant	8.75	8.38	2.14	2.07	8.28	1.15	0.40	4.08	-2.97	-0.36
	12.67	11.77	6.16	5.85	10.29	3.23	0.75	7.91	-2.53	-0.38
Observations	2611	2496	2611	2496	450	2161	771	1017	416	439
R-squared	0.68	0.69	0.88	0.88	0.93	0.88	0.95	0.83	0.75	0.76
RMSE	0.09	0.09	0.17	0.17	0.14	0.15	0.13	0.16	0.25	0.23
	t statistics below coefficient estimates									

Table V-2: Traits of Twelve Counties where the Reduced Form Fits Well															
County Name	MSA Name	State	HP_CA2	HP Growth 2000-05	HP Growth 2005-10	Annual Sales	Rent to Value	Value/ Price	Foreclosure Inventory	HH Income	Share of High Cost Loans in 2004-05	Fixed Effect	Residual	IQR
Allegheny County	Pittsburgh	PA	$ 110,908	28%	-9.1%	9,905	5.4%	2.9	1.4%	$ 47,505	14%	3%	3%	4%
Baltimore County	Baltimore-Towson	MD	$ 232,101	81%	-7.1%	5,700	3.1%	4.3	1.3%	$ 62,543	20%	-1%	-3%	4%
Beaver County	Pittsburgh	PA	$ 87,084	28%	0.4%	1,477	4.9%	2.8	1.1%	$ 43,898	19%	-3%	-4%	4%
Dauphin County	Harrisburg-Carlisle	PA	$ 135,130	30%	0.4%	2,534	5.3%	2.9	1.3%	$ 52,177	15%	0%	-6%	7%
Madison County	St. Louis	IL	$ 104,092	37%	-0.4%	2,441	6.7%	2.7	2.0%	$ 51,901	16%	1%	-8%	9%
Marion County	Salem	OR	$ 181,216	28%	-0.5%	2,637	4.2%	4.5	1.5%	$ 44,570	16%	4%	3%	5%
Medina County	Cleveland-Elyria-Mentor	OH	$ 156,894	20%	-13.1%	1,461	5.8%	2.5	1.0%	$ 63,543	12%	-4%	-3%	4%
Newport News	Virginia Beach	VA	$ 187,679	76%	3.4%	1,447	4.8%	4.0	1.4%	$ 50,937	20%	-1%	7%	5%
St. Johns County	Jacksonville	FL	$ 231,236	66%	-19.7%	2,602	6.3%	3.2	1.0%	$ 58,888	9%	-4%	0%	2%
Tarrant County	Fort Worth-Arlington	TX	$ 131,664	19%	3.4%	27,400	6.5%	2.5	2.4%	$ 52,385	16%	-1%	6%	4%
Washington County	Hagerstown-Martinsburg	MD	$ 170,490	76%	-28.6%	1,246	4.3%	3.3	2.1%	$ 52,857	18%	0%	0%	6%
Wicomico County	Salisbury	MD	$ 161,967	63%	-16.2%	860	5.4%	3.4	1.9%	$ 49,913	20%	-4%	-9%	2%
Values for Full Sample			$ 221,907	51%	-10.5%	3,625	4.4%	4.3	1.0%	$ 52,966	15%	2%		

Table V-3: Estimates of the Annual Percentage Changes in House Prices (VAR Equation Results)

Dependent Variable	**D.lnHVAL_CA**	**D.lnHVAL_CA**
Exogenous Variables:		
Lagged growth rates in:		
House Prices	**0.25**	0.14
	9.86	5.59
Residual from VEC Regression	**-0.11**	-0.01
	-8.27	-1.01
Median Household Income	**-0.13**	-0.06
	-3.44	-1.68
Annual Rent	0.04	0.04
	1.19	1.10
Volumne of Residential Sales	**0.02**	0.03
	5.80	6.15
Foreclosure Inventory (% of OOwM)	**-5.83**	-6.12
	-12.73	-13.37
Unemployment Rate	0.004	0.00
	2.70	1.96
Share of High Cost Mortgages in 2004-05	**-0.22**	-0.18
	-5.43	-4.12
House Price Growth in 2000-05	**-0.08**	-0.05
	-11.73	-4.62
Share of 2004-05 Employment in Construction	0.18	0.40
	1.65	3.37
SAND State		-0.05
		-6.83
Constant	0.04	0.03
	4.03	1.86
Observations	1733	1733
R-squared	0.37	0.38
RMSE	0.0809	0.0806

t statistics below coefficient estimates. The second column also includes MSA Group fixed effects.

Table V-4: Predictions of House Prices for 2011						
County Name	**MSA_name**	**state**	**House Price**	**Predicted House Price**	**Actual less Predicted HP**	**Rank in Pct Diff**
Outperforming Predictions						
Loudoun County	Washington-Arlington-Alexandria	VA	$ 481,745	$ 286,051	41%	1
Essex County	Newark-Union	NJ	$ 401,054	$ 260,904	35%	2
Williamson County	Nashville-Davidson--Murfreesboro--Franklin	TN	$ 320,054	$ 209,363	35%	3
Fayette County	Atlanta-Sandy Springs-Marietta	GA	$ 213,954	$ 150,185	30%	4
Charles County	Washington-Arlington-Alexandria	MD	$ 280,005	$ 196,683	30%	5
Forsyth County	Atlanta-Sandy Springs-Marietta	GA	$ 202,931	$ 147,153	27%	6
Wright County	Minneapolis-St. Paul-Bloomington	MN	$ 151,161	$ 109,998	27%	7
St. Johns County	Jacksonville	FL	$ 220,476	$ 160,939	27%	8
Delaware County	Columbus	OH	$ 237,765	$ 174,961	26%	9
Dakota County	Minneapolis-St. Paul-Bloomington	MN	$ 214,077	$ 160,706	25%	10
Gwinnett County	Atlanta-Sandy Springs-Marietta	GA	$ 142,269	$ 108,332	24%	11
Placer County	Sacramento-Arden-Arcade-Roseville	CA	$ 286,000	$ 218,858	23%	12
Predictions Match Actual Outcomes						
Tulsa County	Tulsa	OK	$ 131,975	$ 130,428	1%	204
Nueces County	Corpus Christi	TX	$ 141,217	$ 139,698	1%	205
Weber County	Ogden-Clearfield	UT	$ 153,278	$ 151,637	1%	206
Santa Cruz County	Santa Cruz-Watsonville	CA	$ 464,564	$ 459,881	1%	207
Franklin County	St. Louis	MO	$ 118,633	$ 117,472	1%	208
District of Columbia	Washington-Arlington-Alexandria	DC	$ 373,309	$ 369,773	1%	209
Bernalillo County	Albuquerque	NM	$ 178,946	$ 177,468	1%	210
Warren County	Allentown-Bethlehem-Easton	NJ	$ 218,799	$ 217,067	1%	211
Rock Island County	Davenport-Moline-Rock Island	IL	$ 89,892	$ 89,195	1%	212
York County	Charlotte-Gastonia-Rock Hill	SC	$ 136,145	$ 135,298	1%	213
Brown County	Green Bay	WI	$ 129,975	$ 129,174	1%	214
Ocean County	Edison-New Brunswick	NJ	$ 237,796	$ 236,346	1%	215
Winnebago County	Oshkosh-Neenah	WI	$ 129,280	$ 128,506	1%	216
Worst Performers						
Pueblo County	Pueblo	CO	$ 111,996	$ 150,243	-34%	420
Niagara County	Buffalo-Niagara Falls	NY	$ 90,498	$ 121,410	-34%	421
Peoria County	Peoria	IL	$ 104,319	$ 140,600	-35%	422
Clark County	Springfield	OH	$ 74,855	$ 100,951	-35%	423
Beaver County	Pittsburgh	PA	$ 88,812	$ 121,062	-36%	424
Scott County	Minneapolis-St. Paul-Bloomington	MN	$ 190,561	$ 260,669	-37%	425
Denver County	Denver-Aurora-Broomfield	CO	$ 209,010	$ 287,431	-38%	426
Suffolk County	Nassau-Suffolk	NY	$ 342,121	$ 471,461	-38%	427
Lancaster County	Lincoln	NE	$ 131,924	$ 184,177	-40%	428
Westmoreland County	Pittsburgh	PA	$ 90,536	$ 140,419	-55%	429
Linn County	Cedar Rapids	IA	$ 119,248	$ 200,062	-68%	430
St. Louis County	Duluth	MN	$ 110,651	$ 235,083	-112%	431

Table V-5: Drivers of the County Fixed Effects Estimates			
	Ordinary Regression	Stepwise Regression	Stepwise Regression
Dependent Variable	**Fixed effect estimates from Reduced Form**		
Sample	All counties	All counties	All counties with Regulatory Variables
Exogenous Variables			
Distance to Los Angeles	-0.000172	-0.000142	0.00021
	-2.777	-2.329	2.543
Distance to Miami	0.000488	0.000484	0.000377
	11.49	11.44	3.423
Distance to Cook County	0.000334	0.000378	
	5.076	5.868	
Distance to New York City	-0.00056	-0.000577	
	-8.801	-9.012	
County's share of MSA population	-0.179	-0.138	-0.278
	-3.833	-3.001	-2.967
Housing units per square mile	0.00001	0.00002	
	1.085	2.629	
Water per square mile of land area	0.491	0.555	0.25
	4.473	5.067	1.839
Wharton Regulatory Index			0.0799
			1.986
RS Means 2008 MSA Construction Cost Index			0.0080
			5.31
Employment: Construction; (Ths.)	-0.00439		
	-1.799		
Population: Ages 65 and over; (Ths.)	0.00251		0.00141
	3.741		2.682
Personal Bankruptcies	(0.00003)		(0.000020)
	(3.32)		(2.39)
Net Migration	0.00717	0.00664	0.00651
	2.444	2.54	2.341
Constant	-0.0272	-0.0853	-2.136
	-0.179	-0.568	-9.29
Observations	439	439	216
R-squared	0.474	0.456	0.561
RMSE	0.317	0.322	0.289
t statistics below estimates			

Table V-6: Drivers of the Rent to Value Ratio			
	All years and counties	2005	2010
Dependent Variable	Rent to Value Ratio		
Exogenous Variables			
Median Household Income	**-0.02**	**-0.04**	**-0.04**
	-12.83	-9.75	-9.77
Housing Vacancy Rate	0.00	-0.01	0.03
	-0.77	-0.61	2.16
Regular and REO sales	0.0010	0.0004	0.0037
	3.30	0.42	2.85
Foreclosure Inventory (% of OOwM)	**0.23**	**0.69**	**0.31**
	10.80	4.78	5.61
Property Tax Rate	**1.34**	**0.82**	**0.55**
	15.95	5.70	3.43
Unemployment Rate	0.001	-0.001	0.000
	9.28	-2.44	0.75
One year treasury rate	0.00		
	-0.39		
Ten year treasury rate	0.00		
	-0.71		
CPI Index base 2009	0.00		
	0.19		
Constant	0.26	0.54	0.56
	18.31	11.71	12.06
Observations	2611.00	416.00	439.00
R-squared	0.77	0.49	0.51
RMSE	0.01	0.01	0.02
The measure of R/V used in this regression is the ratio of ACS rent to the CA House price index			
and adjusted by Olsen's measure of constant quality rents (renttoval_olsen_CA).			
Fixed effects for the states and nine MSA groups are included in column 1.			
Fixed effects for the nine MSA groups are included in columns 2 and 3.			

Table V-7: Drivers of the Price to Income Ratio				
	ALL	2005	2010	ALL
Dependent Variable	**Price to Income Ratio (P/Y)**			
Exogenous Variables				
Median Household Income	-1.48	0.71	0.42	0.93
	-10.32	1.31	1.17	5.38
No. of OO HU (OOwM)	**0.92**	**1.38**	**1.25**	**1.28**
	20.75	8.89	9.05	22.68
Housing Vacancy Rate	3.08	3.76	1.36	-2.83
	5.41	1.64	1.05	-3.72
Regular and REO sales	**-0.45**	**-0.41**	**-0.61**	**-0.56**
	-12.58	-3.52	-5.43	-12.45
Foreclosure Inventory (% of OOwM)	**-16.63**	**-113.40**	**-19.90**	**-26.10**
	-6.69	-5.87	-4.19	-8.92
Property Tax Rate	**-269.70**	**-197.30**	**-101.80**	**-147.20**
	-27.95	-10.26	-7.41	-22.45
Unemployment Rate	-0.11	0.21	0.00	0.05
	-8.09	2.92	-0.07	3.16
One year treasury rate	-0.14			0.05
	-4.63			1.19
Ten year treasury rate	0.32			0.28
	4.70			2.91
CPI Index base 2009	-3.82			-9.22
	-4.53			-7.97
Share of Value Owned by HH GE 65				**8.41**
				11.86
Constant	16.68	-13.49	-7.96	-7.71
	10.09	-2.20	-2.00	-3.73
Observations	2611.00	416.00	439.00	2611.00
R-squared	0.72	0.43	0.38	0.44
RMSE	1.11	1.86	1.36	1.56
t statistics below estimates				
Fixed effects for the states and nine MSA groups are included in column 1.				
Fixed effects for the nine MSA groups are included in columns 2, 3, and 4.				

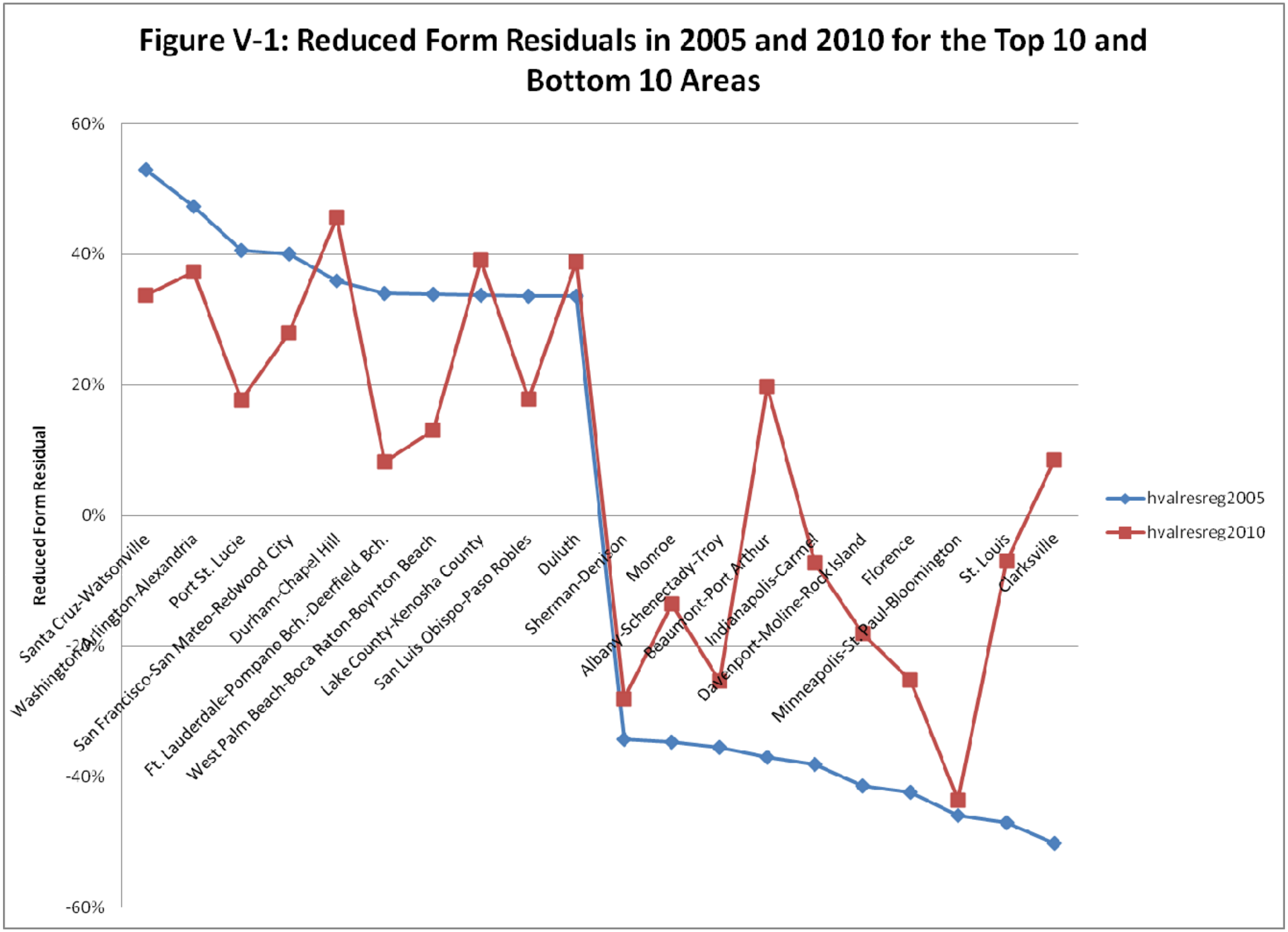


Figure V-1: Reduced Form Residuals in 2005 and 2010 for the Top 10 and Bottom 10 Areas
Reduced Form Residual
60%
40%
20%
0%
-20%
-40%
-60%
Santa Cruz-Watsonville
Washington-Arlington-Alexandria
Port St. Lucie
San Francisco-San Mateo-Redwood City
Durham-Chapel Hill
Ft. Lauderdale-Pompano Bch.-Deerfield Bch.
West Palm Beach-Boca Raton-Boynton Beach
Lake County-Kenosha County
San Luis Obispo-Paso Robles
Duluth
Sherman-Denison
Albany-Schenectady-Troy
Monroe
Beaumont-Port Arthur
Indianapolis-Carmel
Davenport-Moline-Rock Island
Florence
Minneapolis-St. Paul-Bloomington
St. Louis
Clarksville
hvalresreg2005
hvalresreg2010

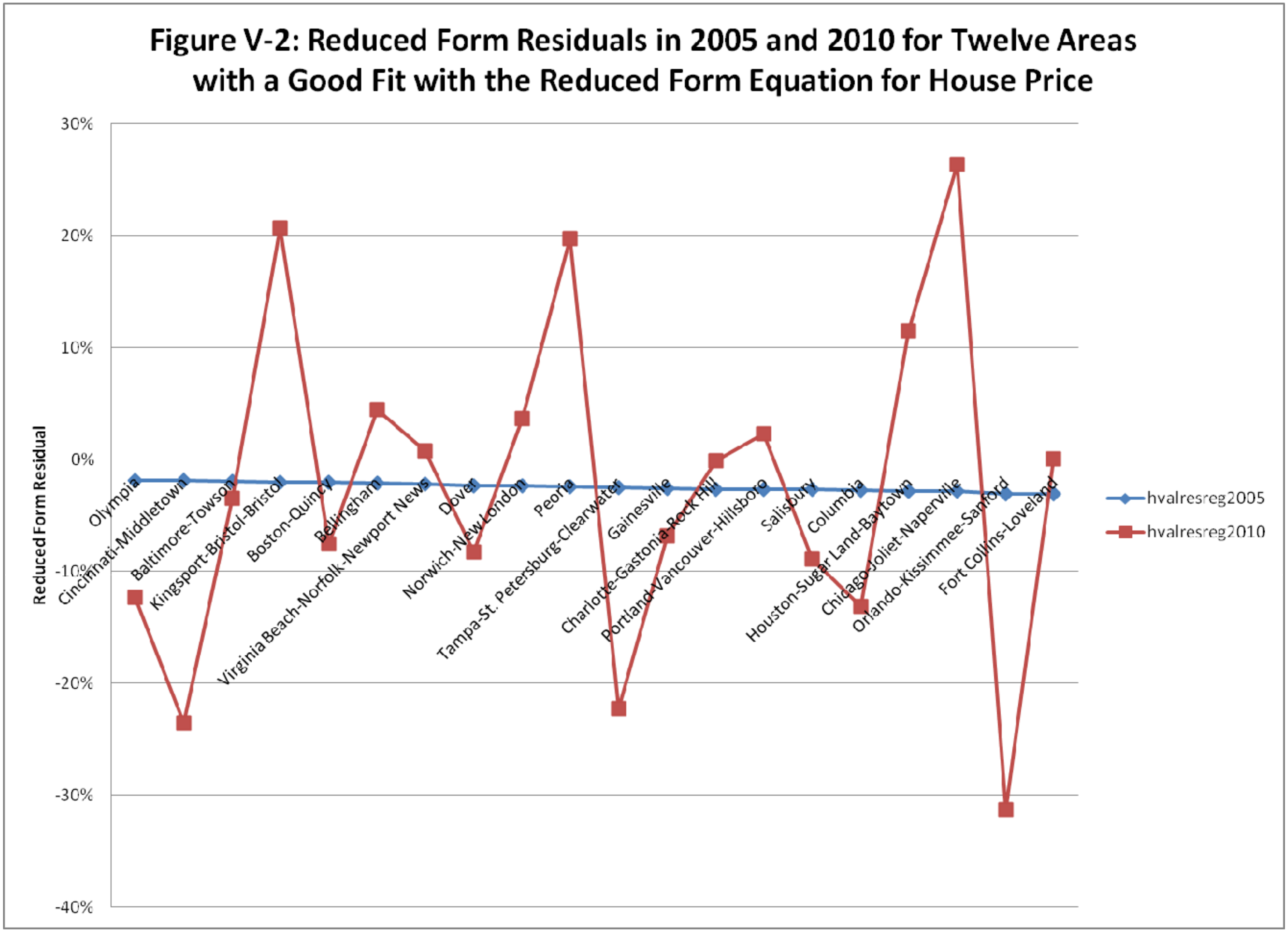


Figure V-2: Reduced Form Residuals in 2005 and 2010 for Twelve Areas with a Good Fit with the Reduced Form Equation for House Price
Reduced Form Residual
30%
20%
10%
0%
-10%
-20%
-30%
-40%
Olympia
Cincinnati-Middletown
Baltimore-Towson
Kingsport-Bristol-Bristol
Boston-Quincy
Bellingham
Virginia Beach-Norfolk-Newport News
Dover
Norwich-New London
Peoria
Tampa-St. Petersburg-Clearwater
Gainesville
Charlotte-Gastonia-Rock Hill
Portland-Vancouver-Hillsboro
Salisbury
Columbia
Houston-Sugar Land-Baytown
Chicago-Joliet-Naperville
Orlando-Kissimmee-Sanford
Fort Collins-Loveland
hvalresreg2005
hvalresreg2010

Figure V-3: Distribution of Fixed Effect Estimates by County

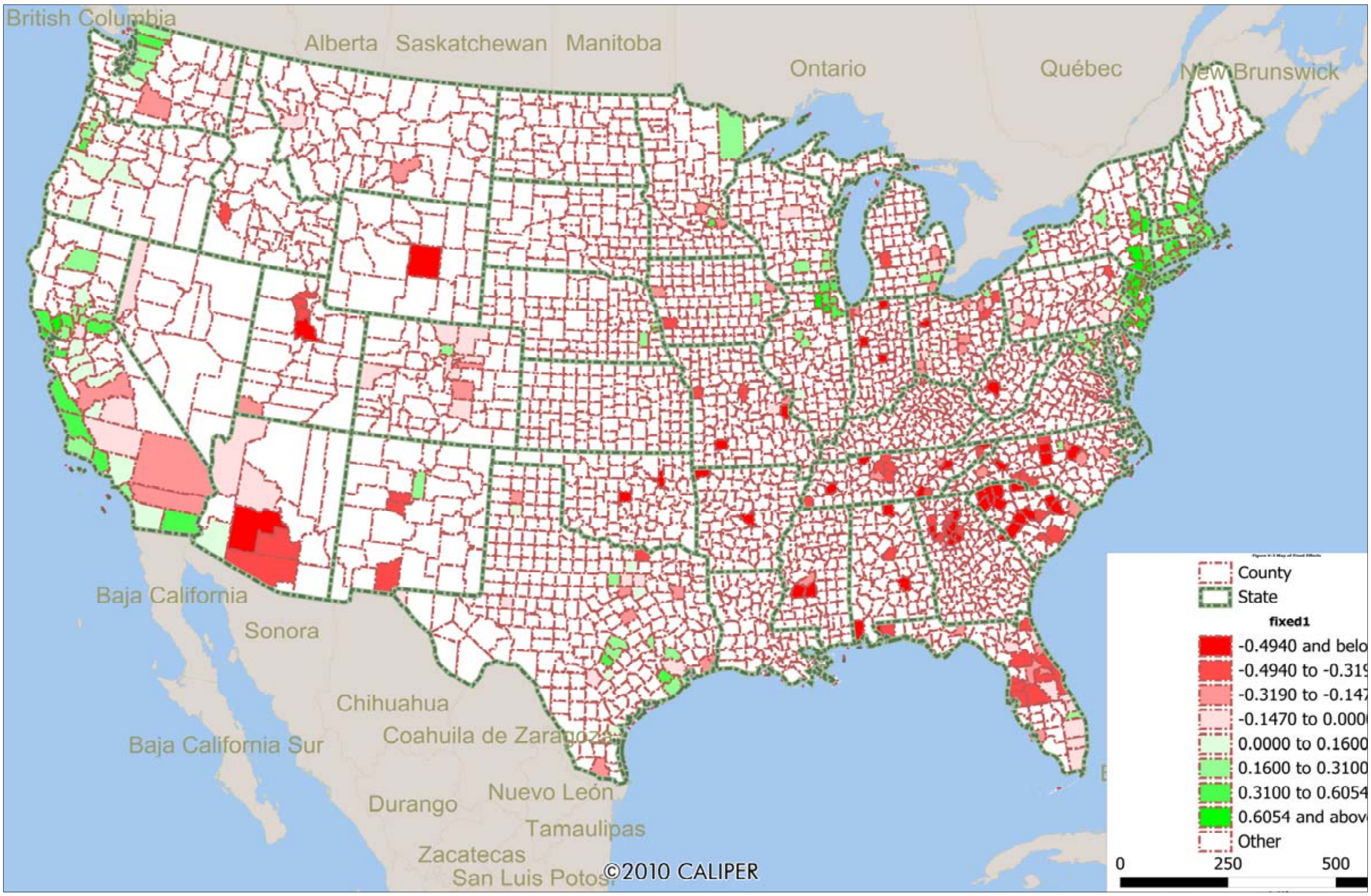


Figure V-4: Distribution of Prediction Errors in House Prices for 2011

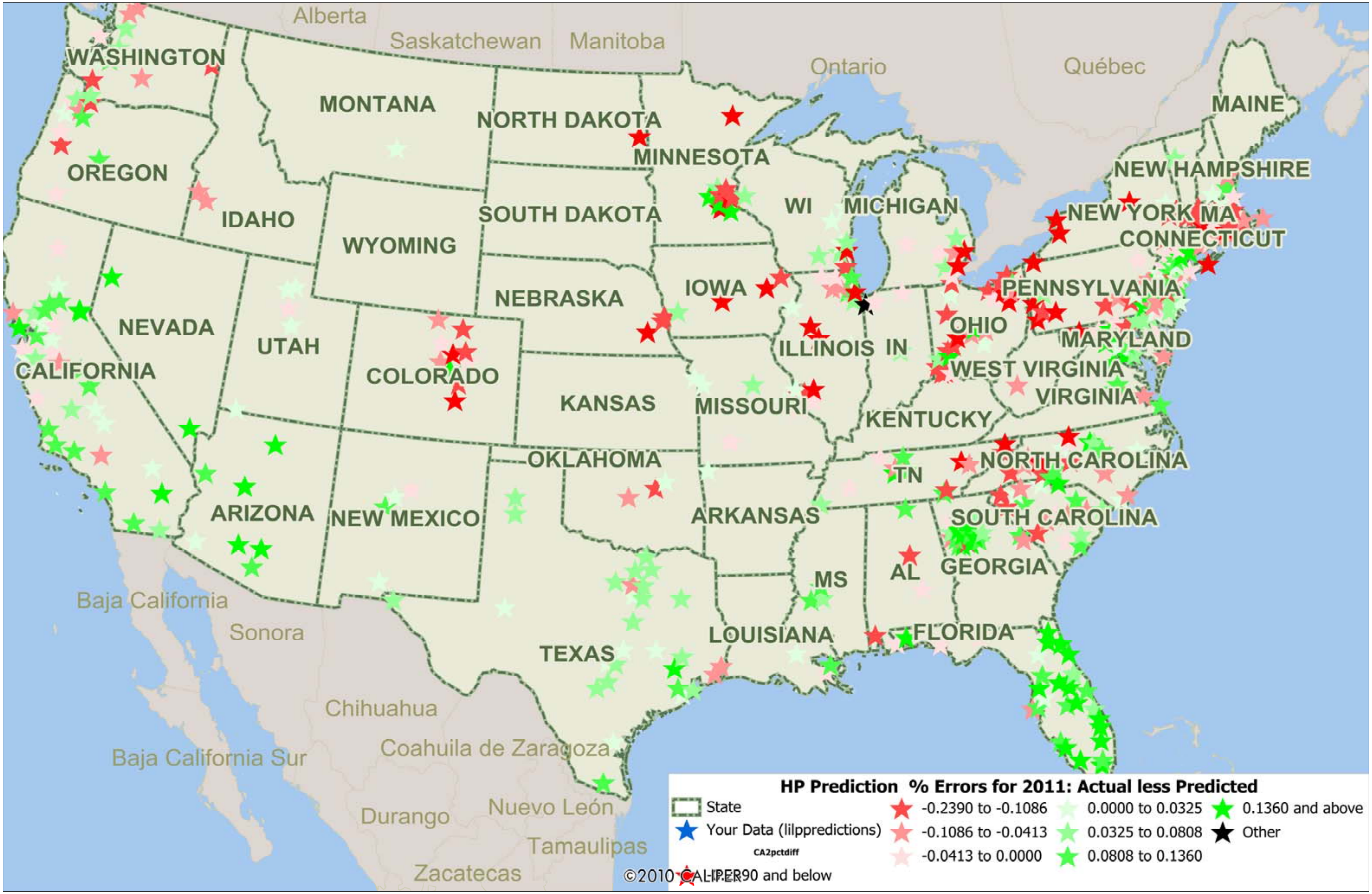


Figure V-5: Comparison of Actual versus Predicted House Prices for 2011

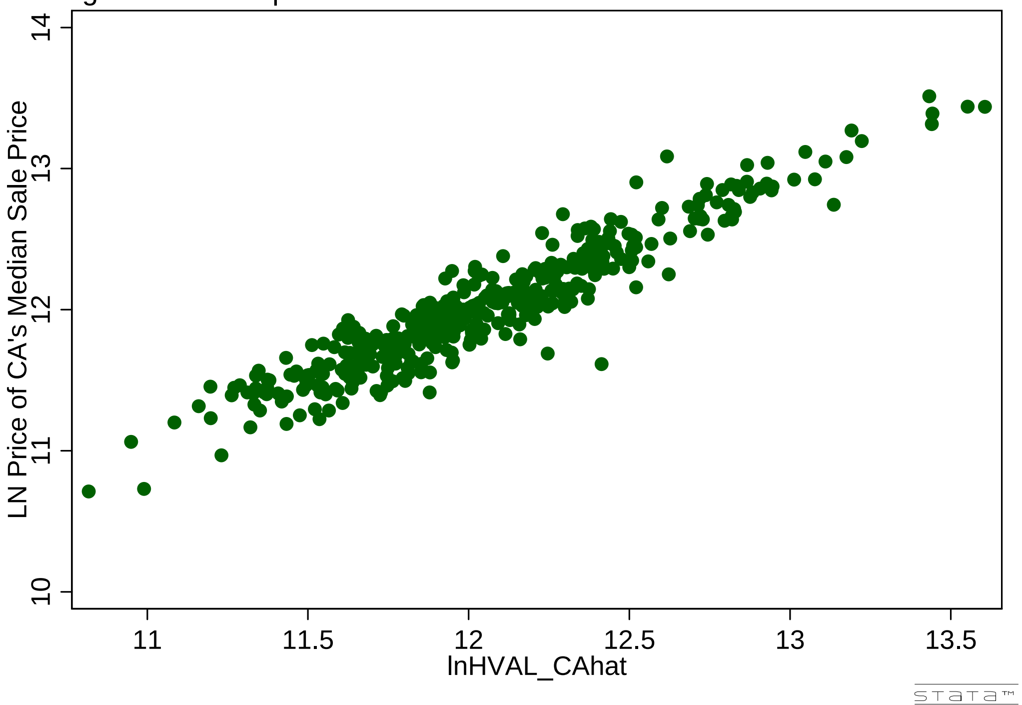
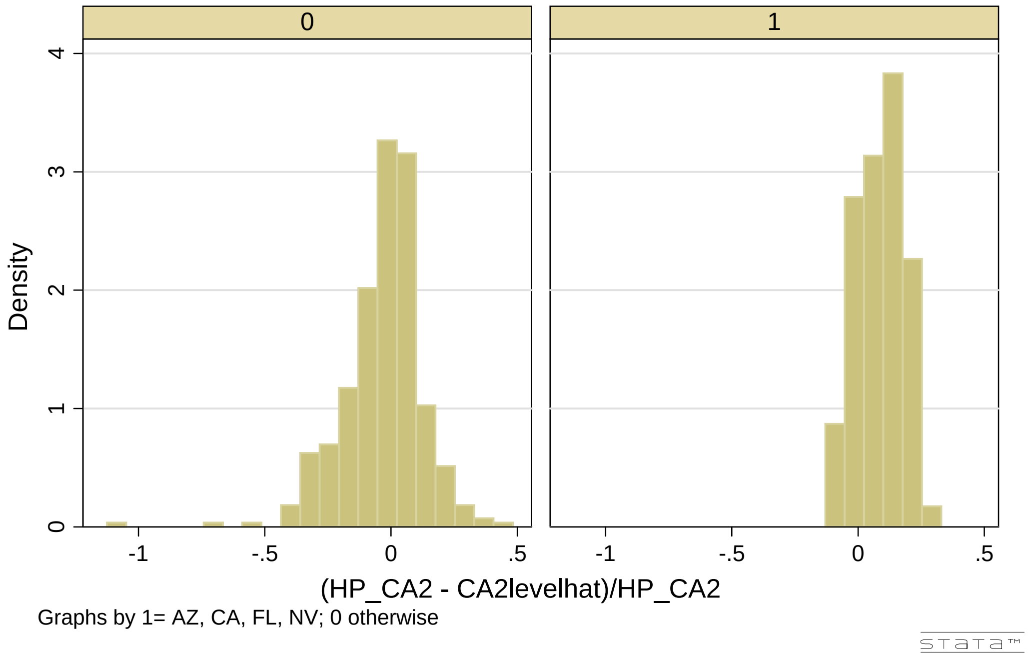


Figure V-6: Distribution of 2011 Forecast Errors of HP (HP - HPhat)
SAND versus non SAND States (percent)



Appendix Table: Statistics on Key Variables						
Variable Name	Label	N	Mean	Std. Dev	Min	Max
B25077_1_EST	OOHU: Median value (dollars)	2611	229680.8	128687.3	61200	922600
HVAL2005	ACS Owner Estimate of Value in 2005	2611	215059.4	130314.7	61200	868200
HVAL_pct	Percentage change in ACS Owner Value since 2005	2611	0.0955921	0.136076	-0.5214543	0.7723005
lnHVAL	Ln of ACS OO Median Value -- B25077_1_EST)	2611	12.22141	0.4772937	11.0219	13.73495
mspsft_annual~j	Price of a 1750 sq ft house with CA data	2611	231248.3	139618.5	55160	1347240
mspsft_ann~2005	Adjusted mspft in 2005	2511	244030.2	150481.6	68192.87	1212470
dmsft_pct_adj	Cumulative pct change in adjusted mspsft since 2005	2511	-0.0075131	0.1683598	-0.6758031	0.7520438
HP_CA2	May 2012 median price series from CA	2437	218366.6	124657.3	50071.07	957899.5
HP_CA2_2005	CA Median Sales price in 2005 with May 15 data	1989	207470.7	115497.1	57837.52	763452.8
lnHVAL_CA2	LN Price of a CAs Median Sale Price received on May 15	2437	12.16591	0.4887943	10.8212	13.7725
dHP_CA2_pct_adj	Annual percentage change in HP_Cas	1989	0.1305973	0.1802508	-0.5094828	1.089363
fhfaall	FHFA HP Price Index value	2611	197.1438	41.56475	118.33	363.58
fhfa2005	FHFA Index in year 2005	2611	190.6775	42.70891	129.53	326.85
dfhfa_pct	Percentage change in FHFA index since 2005	2611	0.0444702	0.1220317	-0.5688311	0.5645455
MSP	Economy.com median sales price in thousands	2611	215443.7	131943	64620	1347240
MSP2005	Median Sales Price of Existing Homes in 2005	2611	226912.5	140542.1	71430	1212470
dMSP2005_pct	Cumulative Annual Pct change in Median Sales Price	2611	-0.0295794	0.1506154	-0.6878531	0.4038231
lnMSP	Ln of Economy.com Median Sales Price	2611	12.14899	0.484782	11.07628	14.11357
lnincome_ACS	Ln of ACS Median Household Income	2611	10.87351	0.2371975	10.10647	11.69141
lnrent_olsen	Ln of Rent Annual Olsen	2611	9.007174	0.1389591	8.517109	9.435905
lnOOwM	Ln of the number of Owner-Occupied Housing Units with at least 1 mortgage	2611	10.87314	0.8574986	9.032648	14.0136
vacancyrate	Total Vacant Units/Total Housing Units	2611	0.1016673	0.0530064	0.0232749	0.4251276
lnallsales_CA	Ln of the Number of Annual Regular and REO Residential Sales	2611	8.159887	1.030658	3.258096	11.9225
fore_cum_pct	Foreclosure inventory as percent of OOwM	2611	0.0093059	0.0125734	0	0.1602856
HOrate	Homeownership rate	2611	0.6861194	0.0948365	0.1875201	0.8927179
proptaxrate	Property Tax Rate on OO Housing	2611	0.0109567	0.005226	0.0017111	0.0292484
FLBRA	Unemployment Rate; (%)	2611	6.398315	2.94235	1.93	29.73
t1	One year Treasury Rate	2611	2.619487	1.809701	0.37	5.01
t10	Ten year Treasury Rate	2611	3.941873	0.6656213	2.74	4.68
cpi_2009	CPI Index base 2009	2611	0.9725902	0.0406594	0.905	1.024

highcostshare	High Cost Loans ($)/All HMDA loans in 2004-05	2611	0.1551207	0.0505726	0.0159543	0.3575669
dfhfa2000_05	Revised measure of the % change in the FHFA price index in 2005	2611	0.4988739	0.308308	0.0898841	1.305984
empconshare	Share of total employment in construction	2611	0.0495918	0.0196139	0.0062614	0.1656744
distance_LA	Distance in miles to Los Angeles	2611	1656.484	687.5349	0.0001189	2705.631
distance_MIAMI	Distance in miles to Miami	2611	1213.747	635.8336	0.000173	5009.171
distance_Cook	Distance in miles to Cook County	2611	816.5761	528.9362	0.0001512	4734.95
distance_NY	Distance in miles to New York City	2611	1002.297	828.5281	0.0000851	5570.075
popshare	Share of county to MSA population	2611	0.4931236	0.3688029	0.0086489	1.002853
stockdensity	Total Housing Units per sqm	2611	536.4503	2096.33	3.120947	37375.61
waterdensity	Share of land area with water	2611	0.0922828	0.1517075	0.0002792	0.7986585
WRLURI	Wharton Residential Land Use Regulation Index (Average)	2186	0.0493766	0.6622792	-1.532346	3.121081
elasticity	Housing Supply Elasticity	2186	2.001897	0.9715367	0.5952661	7.841701
RSMEANS2008	RS Means Construction Cost Index for 2008 by MSA	1365	169769.5	22036.24	137373	224641
allsales_annual	The number of regular and REO sales in year t	2611	4677.383	8011.476	0	150618
dfhfa2000_05	Revised measure of the % change in the FHFA price index in 2005	2611	0.4988739	0.308308	0.0898841	1.305984
annual_sales_~j	CA Sales or FHX1A if sales<=100	2611	5903.365	8328.511	0	150618
saleratio	Ratio of Adjusted sales to OOHU (B25081_1)	2611	0.0574222	0.0340808	0	0.3063187
lnallsales_CA	Ln of All Sales_annual	2611	8.159887	1.030658	3.258096	11.9225
fore_cum_pct	Foreclosure inventory as percent of OOwM	2611	0.0093059	0.0125734	0	0.1602856
rent_annual_ACS	12 times ACS B25064_1 - rent	2611	9923.389	2312.382	5400	19176
rent_annual_o~n	rent_annual_ACS adjusted for quality	2611	8240.614	1151.279	4999.577	12530.29
renttoval	ACS annualized rent to Median House Value	2611	0.0499471	0.0149675	0.016456	0.1057875
renttoval_olsen	Olsen adjustment to ACS annualized rent to Median House Value	2611	0.0446881	0.019412	0.0081781	0.1427259
valtoincome	ACS Home Value to ACS Median Income	2611	4.132566	1.787704	1.634688	14.23224
renttoval_CA	ACS annualized rent to CA Price Index	2611	0.0504441	0.0177826	0.0101007	0.1807832
renttoval_ols~A	Olsen adjustment to ACS annualized rent to CA Price Index	2611	0.0446796	0.0197322	0.0052235	0.1426843
valtoincome_CA	ACS Home Value to ACS Median Income	2611	4.212083	2.083736	1.393519	21.66169
lnHVAL_CA	LN Price of a 1750 sq ft house with CA data	2611	12.22479	0.4757447	10.91799	14.11357
lnrent_olsen	Ln of Rent Annual Olsen	2611	9.007174	0.1389591	8.517109	9.435905
fore_cnt_annual	Number of foreclosure sales in year	2611	964.9318	2619.095	0	42054
reo_cnt_annual	Number of REO sales in year	2611	704.1095	2140.77	0	42561
ne_cnt_annual	Annual number of properties with negative equity	2610	3276.58	7415.873	0	130926

ne_cnt_pct	Inventory with negative equity as percent of OOwM	2609	0.0377116	0.0536151	0	0.4647851
fore_cum	Cum number of foreclosures in post-foreclosure process	2611	903.8188	2486.067	0	44858
fore_cum_pct	Foreclosure inventory as percent of OOwM	2611	0.0093059	0.0125734	0	0.1602856
Dpopgroup	1 slow growing MSAs; 2 = average; 3 = fast	2611	2.094217	0.8223	1	3
msagroup	1 to 9 and combines Dpop and popgroup	2611	6.348908	2.271356	1	9
counthvalca2	Use this to generate bal sample (5 or 6)	2611	5.955956	0.2052332	5	6
fixed1	Fixed effect estimates using all data and XTREG with fe	2611	-2.63E-10	0.4304547	-1.064658	1.32771
hvalres1	Residuals using all data and XTREG with fe	2611	1.04E-10	0.4377708	-1.199521	1.482493
hvalresreg	Residuals using all data and Ordinary Regression	2611	1.85E-10	0.165659	-0.6848964	0.6629792