Bubbles, Post-Crash Dynamics, and the Housing Market

Crocker H. Liu
Cornell University
School of Hotel Administration
Phone: (607) 255-3739
chl62@cornell.edu

Adam Nowak
Department of Economics
University of West Virginia
Phone: (304) 293-7868
adam.d.nowak@gmail.com

Stuart Rosenthal
Maxwell Advisory Board Professor of Economics
Department of Economics and Center for Policy Research
Syracuse University Syracuse, New York, 13244-1020
Phone: (315) 443-3809
ssrosent@maxwell.syr.edu

May 12, 2014

We thank Doreen Longfellow and the Phoenix Board of Realtors for providing us with access to the Phoenix MLS data. We also wish to thank Linda Shaffer at the Maricopa County tax assessor's office and Mary Whelan at Arizona State University for providing us with access to the Phoenix tax assessment data. We are grateful to the Center for Real Estate Theory and Practice at Arizona State University for allowing us to use the ION data. We also thank Sandra Baldwin from Baldwin Luxury Property Marketing (an affiliate of Sotheby's) and Dub Dellis from Walt Danley Realty (an affiliate of Christie's) for helpful discussions regarding the marketing of luxury homes in Arizona. Finally, we are grateful for comments from William Strange, Joseph Williams, and seminar participants at the University of Connecticut, Queens University, Syracuse University, University of Toronto, the New York Federal Reserve Bank, the 2013 Urban Economic Association meetings, and the 2014 Camp Econometrics. All errors are our own.

Abstract

This paper documents and explains previously unrecognized post-crash dynamics following the collapse of a housing bubble. A simple model predicts that speculative developers ensure stable pre-crash relative prices between small and large homes while their post-crash exit allows small-home relative values to fall. Evidence from Phoenix supports the model. Although home prices doubled 2004-2006, relative prices of small-to-large homes remained nearly constant but then plummeted post-crash. As speculative developers return relative prices must return to pre-boom levels, consistent with patterns since 2011. Anticipated mean reversion indicates that cities can reduce post-crash volatility and mispricing by publicizing size-stratified house price indexes.

JEL Codes: G12, G14, R31, R51
Key Words: Housing Bubble, Repeat Sales, Post-Crash Dynamics

I. Introduction

This paper documents and explains previously unrecognized post-crash dynamics following the collapse of a housing price bubble. We develop a simple optimization model which predicts that the presence of speculative developers ensures stable pre-crash relative prices between small and large homes. This occurs because developers direct new construction to the highest yielding market segment. The post-crash exit of speculative developers allows relative prices across home-size segments to diverge while the presence of thick, competitive markets ensures that price divergence will tend to be associated with a greater decline in small home prices relative to large home values.[1] Our model further predicts that as markets recover and speculative developers return, relative prices will return to pre-boom levels. We test the model using 2001-2013 data from Phoenix, Arizona. Results support the model predictions. Moreover, evidence of mean reversion indicates that cities can reduce post-crash volatility and mispricing by publicizing size-stratified house price indexes.

Phoenix, Arizona is a perfect laboratory to illuminate the ideas above and related post-crash dynamics. Phoenix is a large, rapidly growing city surrounded by extensive open, easily developed land. Phoenix also recently experienced one of the most dramatic housing boom-bust cycles ever recorded. Between January, 2004 and January, 2006 housing prices in the Phoenix metropolitan area doubled. This dramatic rise in values is displayed in Figure 1a which plots indexes from a repeat sales model of single family residential housing transactions in Phoenix from January, 2001 to September, 2013.[2] Also apparent in Figure 1a, prices rose at a smooth and modest pace from 2001 through 2003, and then crashed 2007-2009 until stabilizing in early 2009. The intense boom-bust pattern in price is mirrored in home sales in Figure 1b: sales

[1] Details of this argument are provided in Section 5.3 when we discuss the post-crash price dispersion.
[2] The manner in which the repeat sale price index is estimated is explained in Section 3. Ninety-five percent confidence bands are displayed in Figure A-1 in Appendix A and confirm that the index is precisely estimated.

peaked in late 2005 and then fell by over 50 percent by early 2008. Sales have recovered since 2010 while prices have jumped up nearly 50 percent since mid-2011 (Figure 1a).

The patterns in Figures 1a and 1b are well known. Figures 2a and 2b, however, document a further set of patterns that have largely gone unnoticed. Figure 2a plots house price indexes from January, 2001 to September, 2013 for different home-size segments from the 5th to the 95th size percentiles of the market.[3] Notice that prices move together across market segments up to the peak of the boom in 2006. As the crash deepened, however, prices fell below pre-boom levels and small-home prices fell notably further than large-home values.[4] As markets began to recover in 2011, two further patterns are noteworthy. Although construction remained depressed (Figure 2b), it is apparent that relative prices across home-size segments exhibit mean reversion with small-home prices rising further than large-home prices (Figure 2a). These patterns are robust to more refined stratifications of the market as will be presented later in the paper.

It is worth emphasizing that although post-crash exit of spec-home developers allows for the possibility that relative prices across home-size segments could diverge, their exit does not ensure that price divergence will occur. The same is largely true for the other patterns in Figures 2a and 2b highlighted above. To explain these other patterns additional arguments are needed.

---

[3] Stratifying the market in this manner bears some resemblance to recent papers by Leventis (2012), McManus (2013), McMillen (2013), and Landvoigt, Piazzesi, and Schneider (2013), all of which stratify repeat sales measures of housing markets into price tiers (e.g. high, medium, and low). Nevertheless, our emphasis on house size as the stratifying measure is fundamentally different as house size remains fixed across turnover dates in contrast to relative price levels. In addition, Leventis (2012), McManus (2013), and McMillen (2013) focus on statistical properties of the estimated indexes without seeking to explain the underlying patterns. Landvoigt, Piazzesi, and Schneider (2013) do develop a demand side model that matches movers with different levels of wealth, income, and other demographic traits to different price tier homes. Focusing on San Diego 2000-2005, they emphasize that easy access to mortgage credit disproportionately pushed housing demand up among lesser valued homes. They argue that this contributed to upward pressure on prices throughout the market but most prominently in the lower-tier segment. While insightful, their model ignores the supply side of the market and the role of speculative home builders which is a primary focus of this paper. Head et al (2014) develop a model of housing market dynamics that emphasizes search and allows for new construction. However, their model is calibrated to fit aggregate housing market dynamics based on 106 metropolitan areas and does not separate homes into price or size segments.

[4] The dispersion of price indexes between 2009 and 2011 is statistically significant as shown in Figure A-2 of Appendix A which plots the 95 percent confidence bands associated with the different indexes.

As our initial point of departure, we consider the nature of the 2004-2006 price boom as this has implications for post-crash dynamics. Three pieces of evidence suggest that the boom was a bubble driven by unrealistic buyer expectations of future returns. First, despite rapid growth since 1990, the Phoenix MSA is still surrounded by vast amounts of open, easily developed land.[5] This suggests that the long run supply of land for new development should remain highly elastic for years to come, an implication of which is that even large positive demand shocks should have little impact on price.[6] Second, drawing on the present-value model, we evaluate the change in investor expectations of future rent growth necessary to support the doubling of price 2004-2006. We argue that these results strain credibility. Third, quality adjusted sale-to-list price ratios also jumped up in 2004 for small homes up to mansions. We argue that this pattern is consistent with buyer expectations of future housing returns having risen beyond those of seller expectations. This interpretation is especially cogent in the market for mansions. In that sector, the idiosyncratic nature of the homes and related institutional features ensure that sellers never adopt a bidding war marketing strategy in which the list price is set below the anticipated sale price. Based on these arguments we conclude that the 2004-2006 boom was a bubble.

In the aftermath of a price bubble, a simple supply-demand model indicates that price must fall below pre-boom levels in order to clear the market given unwarranted construction in the years prior to the crash. This is consistent with the patterns above. The standard hedonic model of Rosen (1974) further implies that with thick, competitive markets, any divergence in

[5] Haughwout, et al (2012) report that during the 2000-2006 housing boom in Phoenix, quarterly sales of raw land for new residential development typically totaled 10,000-20,000 acres per quarter.

[6] The 1975-2011 history of house price movements in the United States offers a useful comparison. Based on repeat sale indexes from the Federal Housing and Finance Agency (FHFA), Rosenthal (2014), reports that real annual house price growth at the national level in the United States averaged roughly 0.66 percent per year between 1975 and 2011. At the census region level, analogous values range from small negative rates up to a high of 2.2 percent per year for the Pacific region. Measured at these levels of geography, the United States has extensive open land and a very elastic supply function for new development.

relative price across home-size segments will tend to be associated with greater declines in small-home values so that larger homes always sell for smaller homes, *ceteris paribus*. High rates of distressed sales among smaller homes because of job loss and mortgage default likely then pushed small home prices down relative to larger homes, something we provide suggestive evidence of empirically.[7]

As the recovery sets in, our simple model predicts sharp increases in price with limited new construction as observed in Figure 2b. A related implication is that price increases should moderate once price rises to a level sufficient to prompt new construction. At that point, speculative developers will once again become active and relative prices across home-size segments should return to values close to their pre-boom levels. Anticipation of that event helps to explain mean reversion in relative prices which is evident since 2011 (see Figure 2a).

A problem with this modeling structure is that as markets recover, we *expect* small home prices to rise relative to larger home values based on contemporaneous information. Assuming similar risk exposure across home-size segments, this suggests that small homes are expected to yield higher risk-adjusted returns, a violation of the efficient market hypothesis. We believe that the resolution of this inconsistency is that most investors in the Phoenix housing market are *unaware* of the patterns highlighted above. In part, this is because we are unaware of any widely publicized house price indexes for Phoenix or any other major metropolitan area that are stratified based on home-size segments. An implication of anticipated mean reversion is that Phoenix and other cities prone to volatile housing markets can reduce post-crash price volatility

[7] Our findings here are consistent with recent work by Mian, Sufi, and Trebbi (2014). Using state judicial requirements as an instrument for foreclosures, they show that foreclosures contribute to declines in house prices and construction activity. In related work, Gerardi et al (2012) and Haughout et al (2013) show that negative net equity reduces home maintenance expenditures which also contributes to declines in home values.

and mispricing by producing and publicizing size-stratified local house price indexes on a contemporaneous basis.

To develop these and other results, the following section outlines a simple model of the optimization problem faced by a speculative housing developer. Section 3 describes our data and outlines the repeat-estimation method used to produce the various quality adjusted indexes for our analysis. Section 4 provides evidence that the 2004-2006 boom in price was a bubble in the sense that price rose above levels that forward looking investors should have recognized as sustainable. Section 5 examines post-crash house price dynamics and emphasizes the implications of durable housing and turnover. We conclude in Section 6.

II. Speculative developers

This section outlines a simple model that highlights the influence of speculative developers on relative prices across home-size segments of the market. For these purposes, speculative development is said to occur when a developer purchases land and builds a home without a pre-arranged contract with a buyer. This sort of development has been common in Phoenix.

Suppose that housing markets are competitive and there are two types of homes, H_1 and H_2 that generate different levels of housing service flow, h. Spec-home developers purchase N acres of land and then over time construct different combinations of H_1 and H_2 housing on the site to maximize profits treating the land acquisition as a sunk cost. The developer's problem can be thought of as a two-stage maximization problem in which developers first choose the optimal capital-to-land ratios for each housing type and then choose the allocation of land between housing types.

In the first stage, we assume that households maximize welfare by choosing non-housing (x) and housing consumption (h). Households view housing as a composite good comprised of land (L) and capital (k) such that $h = h(L,k)$, where h is twice differentiable and increases with L and k at a decreasing rate. This yields a family of downward sloping, convex indifference curves in L, k space, $g(L,k|h)$, each of which corresponds to a different level of h. Developers produce type-h homes using the cost minimizing combination of L and k at which the marginal rate of substitution of capital for land ($MRS_{k,L}$) equals the factor price ratio $1/r$, where r is the price per unit of capital and the price per unit of land is normalized to 1. Given the modeling assumptions above, this also defines a unique optimal capital-to-land ratio for each type-h home. Alternatively, one could treat local zoning authorities as setting capital-to-land ratios given floor area to lot size ratios (referred to as "FAR") that are common among single family subdivision developments in Phoenix and elsewhere. Regardless, the important feature of the model is that developers are assumed to treat capital-to-land ratios as exogenously given when deciding how to allocate land between housing types in the second stage.

Suppose now that Type-1 housing units are produced using θ_k units of capital and θ_L units of land, while Type-2 homes are produced using 1 unit of capital and 1 unit of land. In the second stage the developer chooses the number of type 1 and 2 homes, N_1 and N_2, respectively, subject to the land supply constraint so as to maximize profits:

$$\begin{aligned} &Max_{N_1 N_2}[P_1 N_1 - r\theta_k k N_1] + [P_2 N_2 - rkN_2] \\ &s.t.\ \ N = \theta_L N_1 + N_2 \end{aligned} \tag{2.1}$$

The first order conditions are:

$$\frac{\partial \pi}{\partial N_1} = P_1 - r\theta_k k - \lambda\theta_L = 0 \tag{2.2a}$$

$$\frac{\partial \pi}{\partial N_2} = P_2 - rk - \lambda = 0 \quad . \tag{2.2b}$$

Setting (2.2a) equal to (2.2b), this implies that

$$\frac{P_1}{P_2} = \theta_L - (\theta_L - \theta_k)\frac{rk}{P_2} \quad . \qquad (2.3)$$

To get a feel for the nature of (2.3), notice that rk/P_2 is the structure's share of house value for the reference home which in this case is type-2 housing. Nationwide, that share is often thought to be close to 0.75 but larger in areas where land is cheap and smaller in expensive land markets. As land becomes increasingly cheap, rk/P_2 converges to 1 and relative prices would be driven by the difference in capital across housing types, θ_k. A more general and realistic scenario, however, is that θ_L is close to θ_k. In that case, developers would approximately double lot size if they double interior space, ceteris paribus, which reportedly is characteristic of local FAR ordinances in much of Phoenix and in other urban areas. From (2.3), θ_L close to θ_k implies that the last term in (2.3) will be small and that relative prices will be close to θ_L.[8]

It is worth emphasizing that (2.3) is the outcome of an optimization problem in which developers adjust new supply across housing types so as to equalize marginal profits between type-1 and type-2 housing. While intuitive, this has far reaching implications. It says that when speculative developers are active relative prices across home-size segments should be quite stable even in the face of large increases in home price levels. This fits the pattern of house price movements described earlier for 2001-2006. It is also consistent with simulations based on (2.3) using summary measures on lot size and interior floor space to measure θ_L and θ_k for seven different home-size segments.[9] The stability implied by (2.3) further suggests that any

[8] Note also, that as land becomes increasingly expensive relative to capital, rk/P_2 goes to zero which would also cause relative prices to equal θ_L.

[9] Table 2a reports summary measures for seven home size categories in our data, including median lot size in square feet and median square footage of the floor space. We calculated θ_L and θ_k for each home size segment by dividing median lot size and median floor space by their corresponding values for the 25th-75th home-size percentile which

divergence in relative prices should be temporary, consistent with evidence of mean reversion in relative prices that coincided with the recovery that began in 2011 (see Figure 2a).

An implication of expression (2.3) is that relative prices could diverge if speculative developers are absent. This is especially relevant to the post-crash environment in Phoenix given the patterns in Figure 2b. Housing starts were at 3,000 units per month just prior to the boom, peaked at 6,000 units per month in 2005, and fell back to 3,000 units per month by 2007. By early 2009, housing starts had fallen below 1,000 units per month and have remained below that level through most of 2013. Thus, speculative developers had mostly exited the market by 2009 removing their disciplining effect on relative prices across home-size segments.

Of course, post-crash exit of speculative developers does not ensure a divergence in relative prices. That is because the model above considers only the supply side of the market and only one component of supply (new construction as opposed to existing stock). To explain the post-crash divergence of relative prices described in the Introduction, it is necessary to allow for the effect of durable stocks of existing housing on supply, turnover rates, and housing demand. These features of the market are considered in the sections to follow.

III. Data and repeat-index econometric methodology

3.1 Data

Data are obtained from three sources with details on how to obtain the data outlined in Appendix B. Our primary source is the Arizona Regional Multiple Listing Service (ARMLS).

---

was used as the reference group. Setting the term rk/P_2 to 0.75, we then calculated the impact of a doubling of price levels on relative prices (indicated by a doubling of P_2 in (2.3)), as with the 2004-2006 boom. Relative prices increase roughly 23 percent for both the 1st size percentile and the 99th size percentile of the market, 14 percent for the 1st-5th size percentile, 6.7 percent for the 5th-25th size percentile, fell 4.7 percent for the 75th-95th size percentile, and increase 13 percent for the 95th-99th size percentile. These values are small relative to a doubling of price levels and support the view that speculative developers help to stabilize relative prices across home size segments. This result is also quite robust to reasonable alternative assumptions that allow the cost of construction per square foot to increase with home size.

These data are proprietary and provide information on single family homes listed, sold, or rented from January, 2000 through September, 2013. An important feature of the data is that a home must be listed after January 1, 2000 to appear in the file. For this reason, we report monthly price indexes beginning in January, 2001 as this allows previously listed homes sufficient time to sell and ensures a sufficiently thick sample in the initial year for which our indexes are reported.

From the ARMLS we use data for Maricopa County which includes all of the Phoenix metropolitan area. There are a total of 1,540,593 home listings over this period prior to cleaning the data; after cleaning we have 1,401,801 observations. For homes listed for sale, each listing contains property address and related identifying codes, sale price, days on the market, original and final listing prices, and more. For homes listed for rent the data provide property address and property identifying codes in addition to rent.[10] Numerous attributes of the properties are also provided including date of construction. Floor space was used to stratify the sample into home-size segments. This was done by first forming a sample of all homes that sold at least once during our sample horizon, 2000-2013. Based on that sample we then calculated the size distribution of homes.

A limitation of the ARMLS is that it does not include homes for sale by owner (FSBO) and primarily for that reason does not include the entire market of single family homes. For some of the applications to follow it is valuable to have an accurate count of the annual stock of single family homes present in Phoenix and also the amount of new construction. In addition, it is desirable to confirm the accuracy of house price indexes based on the ARMLS given that it

[10] See the Arizona realtor website http://www.armls.com/ for additional detail on information provided on each listing in the ARMLS. The ARMLS currently has approximately 29,000 realtor subscribers although it is worth noting that not all realtors belong to the MLS system (see http://www.realtor.org/sites/default/files/reports/2012/nar-membership-count-by-state-historical.pdf and http://www.armls.com/about-armls/mission-history). In 2011 the ARMLS was the seventh largest MLS in the nation based on number of members after California, Florida, Texas, New York, New Jersey and Illinois. As of May 25, 2013, Phoenix ranked tenth in terms of the number of total listings in an MLS among all cities in the US (http://www.realtor.com/data-portal/Real-Estate-Statistics.aspx?source=web).

does not include FSBO sales. For both reasons we augment the ARMLS with data from the Maricopa County Assessor's office which includes information on the universe of all single family homes throughout the Phoenix metropolitan area.

Two alternative versions of the assessment authority data were obtained. The first was purchased from ION Data Express (http://www.iondataexpress.com/), a company that repackages and sells assessment data. The ION data report sale price, home size and various other property characteristics for all real estate transactions in Maricopa County. The ION data does not include information on list price, days on the market, or rents for rental units. In addition, while these data were available beginning in 1988 they only extend to November 2011. Between January 2000 and November 2011, the period over which the ION and ARMLS databases overlap, 1,175,566 single family home sales were recorded in the ION data. This compares to 825,186 sales in the ARMLS database. The ARMLS, therefore, contains roughly 70 percent (825,186/1,175,566) of sales registered with the Assessor's office and reported in the ION data.

To validate the price indexes based on the ARMLS data we estimated separate repeat sale price indexes using both the ARMLS and the ION data. This was done for a variety of home-size segments in the market from January, 2001 through November, 2011 and also for both monthly and annual price indexes. Table A-1 in Appendix A reports the correlation between the corresponding price indexes for twenty different home-size segments. The level of correlation is nearly always above 95 percent and usually higher. This confirms that the ARMLS price indexes are reflective of price movements throughout the Phoenix single family home market.

An alternative version of the Maricopa County Assessor office data was obtained through Arizona State University. These data did not contain transaction prices but do allow us to

accurately measure the total number of single family homes present in Maricopa County on an annual basis from 2000 through 2012. This allows us to determine the size of the single family housing stock in a given year by home-size segment in the market. Changes in the amount of housing stock between adjacent years were also used to approximate new construction under the assumption that there are few demolitions in Phoenix during our sample period.

3.2 Repeat index methodology

The empirical analyses to follow evaluate a numbers of series that must first be estimated using the data above. Those series measure the percent change in housing market outcome variables holding constant the quality of the housing units. Outcome measures include home sale prices, sale-to-list price ratios, and rents for rental units. For each index, we use the same methodology as for widely used repeat sale indexes that were first developed by Bailey, Muth, and Nourse (1963) and later popularized by Case and Shiller (1989). More precisely, consider two successive times when a home turns over at time t and $t+\tau$, respectively. For each of these turnovers, the outcome of interest (e.g. sale price) is denoted as Y and can be written as,

$$Y_t = e^{\gamma_t} f(\mathbf{X}_t; \boldsymbol{\beta}_t) \ , \tag{3.1a}$$

$$Y_{t+\tau} = e^{\gamma_{t+\tau}} f(\mathbf{X}_{t+\tau}; \boldsymbol{\beta}_{t+\tau}) \ . \tag{3.1b}$$

where $f(\mathbf{X}; \boldsymbol{\beta})$ is an unknown and possibly non-linear function of the structural and neighborhood characteristics of the home ($\mathbf{X}$) and their corresponding coefficients ($\boldsymbol{\beta}$). The terms γ_t and $\gamma_{t+\tau}$ are the parameters of interest and reflect the difference in Y across home turnover dates, t and $t + \tau$. If $\mathbf{X}$ and $\boldsymbol{\beta}$ are unchanged between t and $t + \tau$, taking logs and rearranging gives the log change in Y between turnovers,

$$\log\left(\frac{Y_{t+\tau}}{Y_t}\right) = \gamma_{t+\tau} - \gamma_t + \omega_{t+\tau} \ , \tag{3.2}$$

where ω is a random error term and $f(\mathbf{X}; \boldsymbol{\beta})$ drops out of the model. For a sample of properties indexed by i ($i = 1, \ldots, n$) that turn over at various dates, the model in (3.2) becomes,

$$\log\left(\frac{Y_{t+\tau,i}}{Y_{t,i}}\right) = \sum_{t=1}^{\tau_i} \gamma_t D_{t,i} + \omega_{t,i} \qquad \text{for home } i = 1, \ldots, n \tag{3.3}$$

where D_t equals -1, 0 or 1 depending on whether a given property at time t turns over for the first time, does not turn over, or turns over for the second time, respectively.

Equation (3.3) is the standard repeat sales specification used by Case and Shiller (1989), Harding et al (2007), and many others.[11] Given the identifying assumptions, time invariant house and neighborhood attributes difference away. The γ parameters can then be estimated by regressing the log change in Y between turnover dates on the vector **D**. Provided that **X** and **β** are unchanged between turnover dates, estimates of the **γ**-vector indicate the rate at which Y changes with the passage of time. By substituting in different outcome measures for Y, we estimate the rate at which house prices change over time, in addition to the rate at which sale-to-list price ratios, and rents also change with the passage of time. We estimate these series both on a monthly basis from January, 2001 to September, 2013 and also on an annual basis 2001-2013. In all cases, we report exponentiated values for the indexes (e^{γ}) with annual index values usually normalized to 100 in 2003 and monthly index values normalized to 100 in January, 2003.

IV. Housing price bubble

4.1 Fundamental value

A long tradition in finance has emphasized that with well-functioning, competitive markets, asset price should equal the discounted stream of expected future rents.

[11] Rosenthal (2014) recently adapted the model to consider changes in the income of arriving occupants across turnover dates. We also extend the model here by applying it to sale-to-list price ratios and rents.

$$P_t = \sum_{t=0}^{\infty} E[R_t] / (1+d)^t \tag{4.1}$$

Nevertheless, direct tests of the present value relationship have proved difficult to defend because of the inherently noisy task of measuring expectations of future returns (e.g. Flood and Hodrick (1990), for example).[12] Bearing this in mind, Figure 3 plots the repeat sale price index as the heavy black line. On the same vertical scale, the repeat rent index is plotted as the black line with lightly shaded squares. Comparing the two series, it is obvious that the rent index is flat in comparison to the house price index. However, when the repeat rent index is scaled by a factor of 10 (the dotted line), it is apparent that price movements lead changes in rent and that the two series approximately mirror each other.

To clarify these patterns, suppose that expected real rents are constant so that (4.1) simplifies to $P = R/d$. Holding d constant, temporal variation in R scaled by $1/d$ should be similar to changes in P. Figure 3, therefore, provides partial support for the present value expression in (4.1) but does not preclude the presence of a price bubble for reasons to follow.

4.2 Expectations of future rent growth

Consider now the following question. By how much would investor expectations of future real annual rent growth have to increase to support a doubling of price from January 2004 to January 2006? To address this question, expression (4.1) is re-written imposing the assumption that real rents grow at a constant annual rate g such that $R_{t+1} = (1+g)R_t$, and that the discount rate remains fixed at d with for $g < d$. For the initial period 0, (4.1) simplifies to,[13]

[12] Shiller (1981) was among the first of the present value studies and showed that stock prices displayed excess volatility relative to dividends. In much of the following literature, a common practice when estimating (4.1) has been to use ex post measures of rent to proxy for ex ante investor expectations, as with Meese and Wallace (1994).
[13] From (4.2), g must be less than d so that $(1+g)/(1/d) < 1$ which is necessary for P_0 to be finite.

$$P_0 = R_0 \sum_{t=0}^{\infty} \left[\frac{1+g}{1+d} \right]^t \quad . \tag{4.2}$$

A doubling of price between periods 0 and k can be written as,

$$2 = P_k / P_0 = \frac{R_k}{R_0} \left[\frac{1+g_k}{1+g_0} \right] \left[\frac{d-g_0}{d-g_k} \right] \quad . \tag{4.3}$$

From Figure 3, the ratio, R_{2006}/R_{2004} based on beginning of year (January) values is 106/96 or 1.1. We impose that value on (4.3) and normalize g_0 to zero so that g_k measures the change in anticipated growth rates as of period k relative to period 0. Solving for g_k then yields:

$$1.82 = \left[\frac{d + dg_k}{d - g_k} \right] \rightarrow g_k = 0.45d - \frac{dg_k}{1.82} \quad . \tag{4.4}$$

For plausible values of d and g_k, the term dg_k is small so $g_k \approx 0.45d$.[14]

Expression (4.4) implicitly assumes that investors have an infinite horizon, consistent with an efficient market. In contrast, in Table 1 we highlight the change in g necessary to support a doubling of price between 2004-2006 for investor horizons of infinity, 100 years, 50 years, and 25 years. To simplify, we set R_{2006}/R_{2004} to 1, roughly consistent with the patterns in Figure 3, and normalize the year-2004 value for g_0 to be zero so that g_k reflects the change in g between 2004 and 2006 that would be necessary to support a doubling of price. For each investor horizon, we calculate g_k for a pre-specified real annual discount rate.

The primary challenge in conducting this sort of "what-if" analysis is to select an appropriate discount rate. Some guidance can be gained by considering historical real rates of return on stocks and U.S. government securities. Using data from 1926-2006, Siegel (2008) calculates that the real (constant dollar) annualized return on stocks was 6.8 percent per year

[14] For example, with $d = 0.02$ and $g = 0.01$, dg is 0.0002 and can be ignored in expression (4.4).

based on geometric averaging. For 10-year US Treasury bonds Siegel reports an analogous return of 2.4 percent and for 3-month Treasury bills 0.7 percent (see Siegel (2008), pages 13 and 15). We believe that most investors would perceive real estate to be a riskier investment than short and long term U.S. government securities but safer than a balanced portfolio of stocks. This suggests that a credible real discount rate for real estate investments could be in the neighborhood of 3 percent. In Table 1 we experiment with discounts rates from 2.0 to 5.0 percent in 0.5 percentage point increments.

In Table 1, observe that with an infinite investor horizon, the change in g necessary to support a doubling of price between 2004 and 2006 varies from 1 to 2.5 percentage points for discount rates from 2 to 5 percent, respectively. Even with a 2 percent discount rate, which strikes us as conservative, g_k is quite large when one considers the compounding effect on real rents over an infinite horizon. For shorter horizons, as with 50 or 25 years, the values for g_k are even larger. As an example, with a 3 percent real discount rate, g_k equals 3.12 percentage points for a 50 year horizon and 5.53 percentage points for a 25 year horizon. At these rates, real rents in Phoenix twenty-five years in the future would be 2.15 and 3.84 times higher than at present, respectively. Even if one assumes a very low real discount rate and a very long investor horizon, the values for g_k in Table 1 would require an extraordinary growth in population and employment to be realized given the extensive degree of developable land surrounding Phoenix.

4.3 Sale-to-list price ratios

Although the analysis above is suggestive that the 2004-2006 price boom was a bubble, it suffers from the limitation that we do not actually know the rate at which anticipated returns should be discounted. This section adopts a different modeling strategy that avoids that problem.

With efficient markets, all investors should have access to the same publicly available information and should have similar expectations of future rents in expression (4.1), on average. This implies that any difference between seller versus buyer assessment of a property's current market value should be small and stable over time.[15] Figure 4 provides evidence on this point. Figure 4 displays repeat indexes for the sale-to-list price ratio based on the original list price and also the final list price with the indexes normalized to 100 in January, 2003 in both cases.[16] It is worth emphasizing that the sale-to-list price indexes indicate the percentage change in sale-to-list ratios over time but do not say anything about the level of those ratios. For reference, the sale price index is also provided and all three indexes are plotted from January, 2001 through December of 2006 in order to highlight patterns leading up to and during the boom.[17]

Notice the sharp run-up in sale-to-list price ratios in 2004 regardless of whether one uses the original list price or final list price. Also apparent, in 2004 the sale-to-list index based on original list price is above the index based on final list price while the reverse is true after 2005. This indicates that in 2004 sellers increasingly revised their list price *upwards* while after 2005 there was an increasing tendency for sellers to revise list price down.[18]

Two competing views of how sellers set list price lead to quite different interpretations of these patterns. Under the "conventional" view, sellers set list price above anticipated selling

---

[15] Although we have no particular reason to think that buyers and sellers have different discount rates, our arguments below only require that any difference in buyer/seller discount rates is stable over time.

[16] The indexes were estimated as described in section 3 with the sale-to-list price ratio for a given home turnover substituted for Y_t in expression (3.3).

[17] Both sale-to-list price indexes plummet after 2007 and are quite volatile. This presumably reflects that as the crash deepened sellers initially resisted setting their list prices low and the subsequently cut their asking price as their homes did not sell. This is consistent with evidence of possible loss aversion and concerns about regaining lost equity in Genesove and Mayer (2001) and Engelhardt (2003). Downward price stickiness could also occur if sellers are risk-neutral housing wealth maximizers as in Merlo, Ortalo-Magné, and Rust (2008).

[18] While most revisions to list prices occur when sellers cut their asking price, direct comparison of original and final list prices for individual homes confirms that in 2004 there was an increasing tendency for sellers to revise their asking prices upwards.

price by an amount that increases with uncertainty about a home's market value.[19] Figure 5 presents evidence consistent with that view. The figure plots the median sale-to-final list price ratio for homes in three size segments (based on square footage of floor space), the 25th-75th percentile, the 95-99th percentile, and above the 99th percentile. The median floor space for these home-size segments was 2,048, 4,623, and 6,784, square feet respectively (see Table 2a) so these latter two categories are very large homes that trade in thin markets in which there is considerable uncertainty about a given home's market value. For homes in the 25-75th size percentile, the median sale-to-list price ratio is roughly 98 percent over the 2001-2003 pre-boom period when markets were stable. The median sale-to-list price ratios are always notably lower for homes in the 95th to 99th size percentile and lower still for mansions above the 99th size percentile. These patterns reinforce the view that sellers tend to adopt a list-high-sell-low marketing strategy as uncertainty increases about market value.[20] Under that view the 2004 run-up in sale-to-list ratios implies that seller and buyer expectations of future returns diverged so that sellers were increasingly surprised by unexpectedly high bids. The increasing tendency for sellers to revise upwards their asking price in 2004 (see Figure 4) further reinforces that interpretation. Such divergence in seller-buyer expectations of returns, however, should not occur with informed, forward looking agents.

[19] This view is reinforced by seller agent contract provisions that typically oblige sellers to compensate the agent if a bid is received without contingencies at or above asking price irrespective of whether a sale occurs. Han and Strange (2013) note this feature of the seller-agent contract as well in explaining why the list price is relevant. Partly for these reasons, the seller's list price is often perceived as an implicit promise to sell the home if a bid comes in at or above asking price.

[20] Sass (1988), Glower, Haurin and Hendershott (1998), and Haurin, et al (2010) all find that sellers of unusual homes set higher list prices relative to sale price in comparison to homes with more common attributes and which sell in thicker markets. Additional support for this view is provided by Zuehlke (1987) and Haurin (1988) who show that common style homes and those sold in thicker markets sell more quickly, consistent with lower asking prices. See also Yavas and Yang (1995), Anglin, Rutherford and Springer (2003), Merlo and Ortalo-Magne (2004), and Allen, Rutherford and Thomson (2009) among others for related evidence.

A completely different view of how homes are sold is that as markets heat up, sellers increasingly adopt a "bidding war" marketing strategy in which they intentionally set list price below the anticipated selling price. The intent in such a strategy is to generate excitement, simultaneous bids, and a pseudo auction that results in a quicker sale at a higher price (e.g. Han and Strange (2013, 2014), Haurin et al (2013)).[21] If sellers shifted towards such a strategy as the boom developed that could account for the 2004 spike in sale-to-list ratios in Figure 4.

The market for mansions provides an opportunity to control for confounding effects of seller marketing strategies. For a variety of conceptual and institutional reasons, sellers of mansions never adopt a bidding war strategy. Mansions have highly idiosyncratic features and sell in thin markets where sellers face considerable uncertainty about market value.[22] This greatly increases the risk that multiple bids on a mansion would not arrive at the same time even with list price set well below expected sale price (see Han and Strange (2013) for related discussion). Moreover, because mansions require considerable effort to market seller-agent contracts include added provisions to ensure that the agent is compensated if a seller withdraws after a contingent-free bid at or above list price is received.[23] This suggests that sellers of

---

[21] Han and Strange (2013, 2014) provide the most careful assessment of the frequency, nature, and determinants of bidding wars. At the national level, they show that in the 1980s and 1990s, bidding wars – proxied by instances in which homes sold above list price – accounted for roughly 3-5 percent of sales. That number grew to 20 percent in 2000 and has since dropped back to roughly 10 percent. Han and Strange (2014) further show that bidding wars are more common in thick markets with many potential home buyers and large numbers of sales. Haurin et al (2013) recently examined the Columbus, Ohio housing market and concluded that seller marketing strategy must be based at least in part on an auction-like process. See Horowitz (1992), Chen and Rosenthal (1996a, 1996b), Arnold (1999), Haurin et al. (2010), and Carrillo (forthcoming) for related discussion.

[22] A Walt Danley newsletter (http://waltdanley.com/blog/why-zillow-com-zestimates-are-wrong-luxury-homes/) explains why automated valuation models as used by Zillow.com do not work well for luxury homes. Danley emphasizes that luxury homes have many idiosyncratic features and trade in thin markets. Buyers of mansions may also view the list price as a signal of the home's quality and be deterred by a low list price. This sentiment is echoed in Donald Trump's book, The Art of the Deal (page 122): "The sort of wealthy people we were competing for don't look for bargains in apartments. They may want bargains in everything else, but when it comes to a home, they want the best, not the best buy. By pricing its apartments lower than ours, Museum Tower had just announced that it was not as good as Trump Tower."

[23] The duration of the seller-agent contract for mansions is longer (at least one year) than typical contracts as mansions take longer to sell and are often advertised well beyond the immediate metropolitan area. Mansions are shown only by appointment and potential buyers must first go through a prequalification process in which their

mansions who set low asking prices risk receiving a single offer at list price and being forced into selling their home at a discount. For these reasons, sellers of mansions virtually *never* intentionally set list price low relative to expected value.[24]

Figure 6a presents annual repeat sale price indexes for the core of the market (25-75th size percentile) and mansion segments of the market (95-99th size percentile and > 99th percentile). The plots confirm that mansion prices boomed along with the rest of the market. Moreover, Figure 6b shows that mansion repeat sale-to-list price ratios display the same jump in 2004 as the rest of the market.[25] Given that mansions are not marketed through a bidding war strategy, we view this as evidence that buyer-seller assessments of market value in the mansion market increasingly diverged in 2004, consistent with mispricing and market inefficiency. Along with the other evidence previously discussed, it seems likely that the Phoenix housing market experienced a bubble over the period driven by unrealistic buyer expectations of future returns.[26]

V. Post-crash price dynamics

5.1 Overview

Section 2 highlighted the influence of speculative developers on relative prices across home-size segments. The model yielded the important result that relative prices should be largely unchanged when the housing market is growing and speculative developers are building

---

financial status is documented before an invitation to view the property is extended. Upon placing a bid on the home, prospective buyers of mansions are obliged to put down "earnest money" of roughly 3-5 percent of the asking price. Additional earnest money must be put down after the appraisal and loan approval are completed. The broker is entitled to 50 percent of the earnest money or commission specified in the contract, whichever is less if the deal fails to occur.

[24] It is also worth noting that prior to 2004 about 10 percent of homes in the core of the market (25-75th size percentile) sold above original list price while for mansions that number was about 5 percent. In 2004, the share of homes that sold above list jumped to roughly 45 percent in the core of the market; for mansions the number increased to only about 7 percent.

[25] 95 percent confidence bands for the plots in Figures 6a and 6b are presented in Figure A-3. Although estimates for the top 1 percent home-size segment are clearly noisy, the patterns described above are still evident.

[26] Foote, Gerardi, and Willen (2012) also argue that homebuyers acted on overly optimistic beliefs regarding house prices during the recent housing boom.

new homes. As noted earlier, that result is consistent with pre-crash patterns in Figure 2a for 2001-2006. The model also implies that any post-crash divergence in relative prices should disappear as markets begin to recover and speculative developers return to the market, consistent with 2011-2013 patterns in Figure 2a. Nevertheless, the post-crash exit of speculative developers is not sufficient to ensure other patterns in Figure 2a that were highlighted earlier. Those patterns include: (i) post-crash prices fell below pre-boom levels, (ii) post-crash prices fell notably further for smaller-home market segments, and (iii) between 2011-2013 construction remained depressed even though prices jumped roughly fifty percent. To explain these patterns additional arguments are required.

5.2 Magnitude of the post-crash decline in price

We begin by assuming that over the pre-boom 2001-2003 period prices were determined solely by the underlying fundamentals of supply and demand: no bubble was present at this time. In Figure 7, the pre-boom price corresponds to P^E at the intersection of supply and demand. Given evidence from the previous section, we assume that the 2004-2006 boom was a bubble in the sense that price rose above levels that could be sustained given underlying fundamentals of demand as might have been anticipated by forward looking investors. In Figure 7, this is illustrated by an increase in price to P^B. That leads to an expansion of the housing stock from H^E to H^B as developers build additional housing but without sustainable demand support.[27] When the bubble bursts as in 2007, and with durable housing (e.g. Glaeser and Gyourko (2005), Haughout et al (2012)), price falls back to P^C on the demand curve where markets clear. This corresponds to the leveling off of price in 2009 in Figure 2a.

[27] It is noteworthy that housing starts peaked in 2005 roughly one year before the peak in home prices (see Figures 1a and 2b). This is suggestive that forward looking speculative developers had begun to recognize increased risk of a market crash prior to the 2006 peak in price.

An implication of the model in Figure 7 is that with durable housing (e.g. Glaeser and Gyourko (2005) and Houghout et al (2012)) and downward sloping market demand, price must fall below pre-boom levels following the collapse of a bubble. This is consistent with the patterns noted in Figures 1a and 2a. In Figure 7 it is equivalent to the condition that P^C-P^E < 0. To further verify this prediction, we stratify the housing market into twenty home-size segments based on the square footage of the floor space in a home for the 0-5 percentile, 5-10 percentile, … and 95-100 percentile.[28] Figure 8 plots annual repeat sale indexes for each home size segment over the period 2001-2013. For each of these indexes, Table 3 tabulates corresponding price index levels by home size segment for the pre-boom index level (P^E) in 2003, the bubble price (P^B) in 2006, and the post-crash price (P^C) in 2009. Also displayed in Table 3 are the differences P^B-P^E, P^B-P^C, and P^C-P^E. Several patterns in Figure 8 and Table 3 are noteworthy. The first is that except for the very largest homes, post-crash price index levels fall well below pre-boom levels. This is consistent with the arguments above and provides further support for the view that the price boom was a bubble.

5.3 Post-crash price divergence across home-size segments

A second pattern in Figure 8 and the second-to-last column of Table 3 (P^C-P^E) is the strikingly monotonic ordering of the difference between the pre-boom and post-crash price index levels across home-size segments. Those differences include -47 for the 0-5th percentile, -10 for the 50th-55th percentile, and 9 for the 95th-100 size percentile. This monotonic pattern of price index divergence is not a coincidence. Instead, it ensures that larger homes always sell for a

[28] Tables 2a and 2b provide summary measures for seven size categories that span the market with medians in Table 2a and means in Table 2b. For reference, observe that median home size is roughly 1,000 square feet for the bottom 1st percentile, 2,000 square feet for the 25th to 75th percentiles, and 7,000 square feet for the 99th percentile.

higher price, *ceteris paribus*, as must be the case with thick, competitive markets and convex preferences.

To clarify, suppose that all homes are identical except for size and households differ only in their willingness to pay for different size homes. With competitive markets each home sells to the highest bidder so that in equilibrium households that place greater value on living space occupy larger homes while those with a lesser willingness to pay occupy smaller homes (as in Rosen (1974) and recent papers by Landvoigt, Piazzesi, and Schneider (2013) and Guerrieri, Hartley, and Hurst (2013)). With sufficiently thick markets, the distribution of home-size types is approximately continuous so that sale prices of small to large homes trace out a smooth hedonic price function, the slope of which defines the price of additional living space. Households then secure a level of housing h^* that equates the price of living space with their marginal rate of substitution between housing and non-housing consumption. With convex preferences, a household would always be better off if it could occupy slightly more housing than h^* at no additional cost, but competition from families with slightly stronger taste for living space precludes that possibility. This ensures that larger homes must sell for a higher price than smaller homes, *ceteris paribus*. This also explains why in a thick market such as Phoenix with a near continuous size distribution of homes, any post-crash divergence will tend to be associated with a greater decline in smaller home price index levels. What this argument does not explain is why price divergence occurred in the first place. We focus on that question next.

Given the similarity of the price boom (P^B-P^E) across market segments, one cannot appeal to the size of the bubble to explain the monotonic ordering of P^C-P^E. A different possibility is that long run supply is more inelastic for larger home-size segments as that would imply less overbuilding in response to the price boom and a smaller spread between P^C and P^E.

It is true that mansions are only built in select neighborhoods and in principle that could contribute to inelasticity of new supply at the very top of the market. Nevertheless, differences in long run elasticity of supply across home-size segments do not seem to offer a plausible explanation for the market-wide monotonic pattern evident in the second to last column of Table 3. The technology to build a 2,000 square foot home, for example, is essentially identical to that used to build 1,500 and 2,500 square foot homes.

Two alternative mechanisms seem more likely to account for the monotonic increase in post-crash drop in price among smaller home segments. The first is that demand for smaller homes could be more inelastic than demand for larger units which, from Figure 7, would increase the magnitude of the price crash. However, we have no way of testing whether demand for smaller units is more inelastic.[29]

A different mechanism is that post-crash job loss and related mortgage default may have contributed to large numbers of distressed home sales that hit small-home market segments especially hard (e.g. Gyourko and Tracy (2014)). This would have contributed to an increase in the share of housing stock listed for sale and related downward pressure on prices.[30] Evidence of this possibility is confirmed in Table 4a. Notice that the share of sales classified as "distressed" in the ARMLS was roughly 1 percent per year from 2000-2006 but then jumps to a peak of 46 percent in 2009. In Table 4b it is evident that between 2007 and 2012 distressed sales were in fact much more common in smaller home market segments. If speculative developers had remained active following the crash, expression (2.3) would have prevailed and the spike in

[29] Occupants of large homes have the financial means to occupy smaller dwellings but the reverse is less true for occupants of smaller homes. In addition, the extra square footage associated with a very large home is a luxury purchase. For both reasons, demand for smaller homes could be inelastic relative to demand for larger homes.
[30] Mian et al (2014) provide evidence consistent with the view that distressed sales push housing prices down while Head et al (2014) suggest that turnover contributes to housing market dynamics. Both studies in part motivate our focus on list-to-stock ratios and distressed sales in the discussion below.

mortgage default along with related shifts in the share of stock listed for sale should not have caused relative prices to diverge across home-size segments. With speculative home building absent, however, that discipline is removed which allows for the possibility that price divergence could emerge. Table 5 presents regressions that provide evidence on this point.

Table 5 presents separate regressions for the pre- and post-crash period (2001-2006 and 2007-2012, respectively) treating the 20 home-size segments as an annual panel. The regressions are of the following general form,

$$\frac{P_{t,s}}{P_{t,50-55^{th}\,Pctl}} = d_t + \theta_1 Log(NewlyBuilt)_{t,s}$$

$$+ \theta_2 \frac{Listings_{t,s}/Stock_{t,s}}{Listings_{2003,s}/Stock_{2003,s}} \tag{5.1}$$

$$+ \theta_3 Log(NewlyBuilt)_{t,s} \frac{Listings_{t,s}/Stock_{t,s}}{Listings_{2003,s}/Stock_{2003,s}} + e_{t,s}$$

where the dependent variable is the ratio of the price index in segment *s* relative to the price index for the 50-55th size percentile which is used as a base of reference. All of the regressions control for year fixed effects, d_t , so that our focus is on variation in relative prices across home-size segments within a given year. Additional controls include the log number of homes built in each home size segment in year *t*, along with the list-to-stock ratio for segment *s* normalized by its year-2003 value. The first two columns in Table 5 include only the annual fixed effects and the number of homes built in the previous year as controls. The next two columns are based on the full specification in (5.1). The final column in the table adds controls for the ratio of distressed sales to all sales in *s* and *t*. This specification is run only for the post-crash period since there are so few distressed sales prior to 2007.

As anticipated, for the pre-crash period in columns 1 and 3, the model controls are only weakly correlated with relative prices. This is apparent from the within R-square values which are just below 20 percent in columns 1 and 3, and the corresponding joint F-tests of the slope coefficients which suggest weak significance (see the summary measures at the bottom of the table). In contrast, for the post-crash models in columns 2, 4, and 5, the within R-square values are all quite high (0.544, 0.716, and 0.741, respectively). Moreover, list-to-stock ratios display a highly significant negative association with relative prices that tends to go away as builders become more active as indicated by the highly significant positive coefficient on the interaction term. Adding controls for the share of distressed sales in column 5 has little further effect which is suggestive that much of the relationship between distressed sales and relative prices is transmitted through the share of stock listed for sale. Together, these patterns confirm that when speculative builders are active, relative prices across home size segments are largely independent of other market conditions, but that is not true when builders are absent.

5.4 Market recovery

A final implication of Figure 7 pertains to dynamics that arise as markets begin to recover. With a continual increase in population as has occurred in Phoenix, housing demand gradually shifts out and up the inelastic portion of the supply function in Figure 7 along segment BC. This continues until markets return to the elastic portion of the long run supply function at point B. During this time, price should increase sharply (up segment BC) but with little new construction. Only after the market has returned to the long run supply function at point B should price increases moderate and construction levels return to levels more characteristic of the pre-boom period. Figure 2b provides evidence consistent with these predictions in that it

documents a 50 percent increase in price since late 2011 despite continued depressed levels of construction. The plots in Figure 8 and the final column of Table 3 also confirm that relative prices across home-size segments have exhibited a dramatic degree of mean reversion since 2011, with small home prices rising further than larger home values. As a result, much of the price dispersion evident in 2009 has disappeared, presumably as forward looking investors have begun to anticipate the impending return of speculative development to the market.

VI. Conclusion

The boom and bust in U.S. house prices of the last decade triggered a near financial meltdown and the great recession. This also brought new appreciation for the need to better understand house price dynamics and contributed to a host of studies on the cause of housing price booms. This paper extends that literature by documenting and explaining previously unrecognized post-crash dynamics that reveal missing information in the housing market and point to feasible policy measures that can reduce post-crash volatility and mispricing.

Our study draws on single family housing transactions in Phoenix from 2000 to 2013. Based on three different sets of evidence we argue that the doubling of price between 2004 and 2006 was a bubble driven by unrealistic buyer expectations of future returns. We then stratify the market into home-size segments from very small homes up to mansions. Results show that relative prices across home size segments were nearly constant prior to the crash, a pattern we attribute to the presence of speculative developers. With collapse of the bubble, house prices fell below pre-boom levels because of unwarranted construction during the boom. Moreover, with the post-crash exit of speculative developers, increased turnover associated with post-crash job loss pushed small-home values down by up to 60 percent relative to the largest homes.

Regression results confirm the essential role of speculative developers. Results also confirm that as speculative developers return to the market, relative prices revert back to pre-boom levels. The implicit forecast that post-crash small-home investments should yield higher returns than investment in larger homes implies arbitrage opportunities and mispricing of homes. We believe this is possible because post-crash exit of spec-home developers removed an essential disciplining effect from the market and that housing market participants are largely unaware of possible divergence in relative prices across home-size segments. Cities can address this absence of information by publishing size-stratified repeat sale indexes.

The patterns uncovered in this paper likely apply to other growing metropolitan areas that are subject to volatile housing markets. They also suggest that real estate bubbles set off a sequence of dynamic effects that continue well past the initial formation and collapse of the bubble.

References

Allen, M., R. Rutherford and T. Thomson (2009). "Residential Asking Rents and Time on the Market." *Journal of Real Estate Finance and Economics* 38: 351-365.

Anglin, P., R. Rutherford, and T. Springer (2003). "The Trade-off Between the Selling Price of Residential Properties and Time-on-the-Market: The Impact of Price Setting." *Journal of Real Estate Finance and Economics* 26: 95-111.

Arnold, M. (1999). "Search, Bargaining and Optimal Asking Prices." Real Estate Economics 27: 453-481.

Bailey, Martin J., Richard F. Muth, and H. O. Nourse (1963). "A Regression Model for Real Estate Price Index Construction." *Journal of the American Statistical Association* 58, 933-942.

Carrillo, Paul (forthcoming). "To Sell or Not to Sell: Measuring the Heat of the Housing Market." *Real Estate Economics*.

Case, Karl and Robert Shiller J. (1989). "The Efficiency of the Market for Single Family Homes," *American Economic Review*, 79(1): 125-137.

Chen, Yong, and Robert Rosenthal (1996a). "Asking Prices as Commitment Devices." *International Economic Review* 37: 129-55.

Chen, Yong, and Robert Rosenthal (1996b). "On the Use of Ceiling-Price Commitments by Monopolists." *RAND Journal of Economics* 27: 207-20.

Engelhardt, Gary,V (2003). "Nominal loss aversion, housing equity constraints, and household mobility: evidence from the United States." *Journal of Urban Economics* 53: 171-19.

Flood, Robert P. and Robert J. Hodrick (1990). "On Testing for Speculative Bubbles." *The Journal of Economic Perspectives*, 4:2, 85-101.

Foote, Christopher, Kris 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," Public Policy Discussion Paper Series No. 12-2, Federal Reserve Bank of Boston, www.bos.frb.org/economic/ppdp/2012/ppdp1202.pdf

Genesove, David and Lu Han (2012). "Measuring the Thinness of Real Estate Markets." Working paper, Rotman School of Management, University of Toronto.

Genesove, David and Christopher Mayer (2001). "Loss aversion and selling behavior: evidence from the housing market." *Quarterly Journal of Economics* 116: 1233-1260.

Gerardi, Kristopher, Paul Willen, Vincent Yao, and Eric Rosenblatt (2012), "Foreclosure externalities: New evidence," NBER Working Paper No. w18353.

Glaeser, Edward L., Joseph Gyourko, and Albert Saiz (2008). "Housing Supply and Housing Bubbles." *Journal of Urban Economics* 64: 198-217.

Glower, M., Donald R. Haurin and Patrick H. Hendershott (1998). "Selling Price and Selling Time: The Impact of Seller Motivation." *Real Estate Economics* 26: 719–740.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst (2013). "Endogenous Gentrification and Housing Price Dynamics" *Journal of Public Economics*, 100, 45-60.

Gyourko, Joseph and Joeseph Tracy (2014). "Reconciling Theory and Empirics on the Role of Unemployment in Mortgage Default." *Journal of Urban Economics*, 80: 87-96.

Han, Lu and William. Strange (2014). "Bidding Wars for Houses." *Real Estate Economics*, 42(1): 1-32.

Han, Lu and William Strange (2013). "What is the Role of the Asking Price for a House?" working paper.

Harding, John, Stuart S. Rosenthal, and C.F. Sirmans (2007). "Depreciation of Housing Capital, Maintenance, and House Price Inflation: Estimates from a Repeat Sales Model." *Journal of Urban Economics*, 61: 193-217.

Harding, John, Stuart S. Rosenthal, and C.F. Sirmans (2003). "Estimating Bargaining Power in the Market for Existing Homes." *The Review of Economics and Statistics*, 85(1): 178-188.

Haughwout, Adrew, Richard W. Peach, J. Sporn, and Joseph Tracy (2012). "The Supply Side of the Housing Boom and Bust of the 2000s." Federal Reserve Bank of New York *Working Paper* (http://www.nber.org/chapters/c12620.pdf) .

Haughwout, Andrew, Sarah Sutherland, and Joseph Tracy (2013), "Negative Equity and Housing Investment," Federal Reserve Bank of New York Staff Reports, No. 636.

Haurin, Donald (1988). "The Duration of Marketing Time of Residential Housing." *AREUEA Journal* 16: 396-410.

Haurin, Donald., J. Haurin, T. Nadauld and Anthony Sanders (2010). "List Prices, Sale Prices and Marketing Time: An Application to U.S. Housing Markets." *Real Estate Economics* 38: 659-685.

Haurin, Donald, Stanley McGreal, Alastair Adair, Louise Brown, and James R. Webb (2013). "List price and sales prices of residential properties during booms and busts." *Journal of Housing Economics* 22(1): 1–10.

Head, Allen, Huw Lloyd-Ellis, and Hongfei Sun (2014). "Search Liquidity, and the Dynamics of House Prices and Construction." *American Economic Review*, 104(4): 1172-1210.

Horowitz, J. (1992). "The Role of the List Price in Housing Markets: Theory and an Econometric Model." *Journal of Applied Econometrics* 2: 115-129.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2013). "The Housing Market(s) of San Diego." mimeo.

Leventis, Andrew (2012). "Home Price Indexes for Homes in Different Price Tiers: Biases and Corrections," Federal Housing Finance Agency Working Paper 12-1.

McManus, Douglas (2013). "House Price Tiers in Repeat Sales Estimation." Freddie Mac, mimeo.

McMillen, Daniel (2013). "Local Quantile House Price Indices." University of Illinois, mimeo.

Meese, Richard and Nancy Wallace (1994). "Testing the Present Value Relation for Housing Prices: Should I leave My House in San Francisco?" *Journal of Urban Economics*, 35, 245-266.

Merlo, Antonio, Francois Ortalo-Magne, and John Rust (2008). "The home selling problem: Theory and evidence." Working Paper, University of Maryland.

Merlo, Antonio and Francois Ortalo-Magne (2004). "Bargaining over residential real estate: evidence from England." *Journal of Urban Economics*. 56, 192-216.

Mian, Atif, Amir Sufi, and Francesco Trebbi (2014), "Foreclosures, House Prices, and the Real Economy," SSRN, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1722195.

Rosen, Sherwin (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy*, 82:1, 34-55.

Rosenthal, Stuart S. (1999). "Residential Buildings and the Cost of Construction: New Evidence on the Efficiency of the Housing Market," *Review of Economics and Statistics*, 81(2): 288-302.

Rosenthal, Stuart S. (2014). "Are Private Markets and Filtering a Viable Source of Low-Income Housing? Estimates from a "Repeat Income" Model," *American Economic Review*, 104(2): 687-706.

Sass, Timothy (1988). "A Note on Optimal Price Cutting Behavior Under Demand Uncertainty," *Review of Economics and Statistics*, 2, 336-339.

Siegel, Jeremy J. (2008). Stocks for the Long Run, 4th edition, McGraw-Hill, New York, NY.,

Shiller, Robert J. (1981). "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*. 71:3, 421-436.

Trump, Donald J. and T. Schwartz (1988). Trump: The Art of the Deal. Random House, New York.

Yavas, Abdullah and S. Yang (1995). "The Strategic Role of Listing Price in Marketing Real Estate: Theory and Evidence." *Real Estate Economics* 23: 347-368.

Zuehlke, Thomas (1987). "Duration Dependence in the Housing Market," *Review of Economics and Statistics*, 69, 701-709.

Table 1
Increase in Real Annual Rent Growth Rate in Percentage Points Necessary to Support a Doubling in Price from 2004 to 2006[a]

Real Annual Discount Rate	Investor Horizon (Years) Infinite	100	50	25
0.020	0.0100	0.0169	0.0289	0.0534
0.025	0.0125	0.0183	0.0300	0.0543
0.030	0.0150	0.0199	0.0312	0.0553
0.035	0.0175	0.0215	0.0324	0.0564
0.040	0.0200	0.0232	0.0337	0.0574
0.045	0.0225	0.0250	0.0350	0.0585
0.050	0.0250	0.0269	0.0364	0.0596

[a]All values based on annuity expressions with an initial zero rate of anticipated annual rent growth and $R_k/R_0 = 1$ in expression (7).

Table 2a
Median Values for Single Family Detached Homes Sold
July 2000 through September 2013 Based on ARMLS Data[a]

		Home Size Category Based on Square Footage of Homes Sold						
	ALL SALES	< 1%	1 - 5%	5 - 25%	25 - 75%	75 - 95%	95 - 99%	> 99%
Observations	881,628	54,088	78,305	238,474	388,043	102,231	16,790	3,697
Median Sale Price	188,848	105,656	129,503	155,033	217,670	395,620	790,000	1,935,666
Median Final List Price	192,749	107,343	131,473	157,602	222,812	409,948	832,210	2,146,084
Median Original List Price	197,967	109,782	134,012	160,277	228,693	425,263	881,355	2,324,410
Median Year Built	1995	1972	1983	1993	1997	2001	2002	2002
Median Days on Market	42	34	34	36	44	58	81	133
Median Sq Feet (Floor Space)	1,789	1,015	1,224	1,496	2,048	3,223	4,620	6,739
Median Sq Feet (Lot Size)	7,401	6,599	6,541	6,673	7,700	10,642	24,085	44,852
Median Original Sale-to-List	0.870	0.972	0.984	0.985	0.981	0.970	0.955	0.927
Median Final Sale-to-List	0.916	0.985	0.993	0.993	0.990	0.983	0.974	0.955

Table 2b
Mean Values for Single Family Detached Homes Sold
July 2000 through September 2013 Based on ARMLS Data[a]

		Home Size Category Based on Square Footage of Homes Sold						
	ALL SALES	< 1%	1 - 5%	5 - 25%	25 - 75%	75 - 95%	95 - 99%	> 99%
Observations	881,628	54,088	78,305	238,474	388,043	102,231	16,790	3,697
Mean Sale Price	249,176	106,840	132,738	163,126	241,181	460,689	933,867	2,229,300
Mean Final List Price	257,963	108,842	134,961	166,146	247,869	480,403	1,001,989	2,497,100
Mean Original List Price	269,621	113,719	139,976	172,315	257,925	501,679	1,070,680	2,745,835
Mean Year Built	1989	1971	1981	1987	1992	1998	1998	1999
Mean Days on Market	65	53	53	57	67	86	116	173
Mean Sq Feet (Floor Space)	1996	984	1,215	1,485	2,092	3,311	4,740	7,281
Mean Sq Feet (Lot Size)	10,079	7,153	6,984	7,388	9,715	16,917	31,814	58,517
Mean Original Sale-to-List	0.957	0.959	0.967	0.965	0.956	0.941	0.912	0.903
Mean Final Sale-to-List	0.979	0.985	0.987	0.985	0.978	0.967	0.945	0.908

[a]All dollar values are in January, 2010 dollars.

Table 3: Sale Price Index Boom and Bust by Home Size Category
2001 through 2013 Based on ARMLS Data

Home Size Percentile	2003 Pre-Boom Price (P^E)	2006 Bubble Price (P^B)	2009 Post-Crash Price (P^C)	2013 Recovery Price (P^R)	P^B- P^E	P^B- P^C	P^C- P^E	P^R - P^E
0 to 5	100	184	53	87	84	131	-47	-13
5 to 10	100	184	65	100	84	120	-35	0
10 to 15	100	182	69	104	82	113	-31	4
15 to 20	100	183	73	107	83	110	-27	7
20 to 25	100	186	77	111	86	109	-23	11
25 to 30	100	189	79	113	89	109	-21	13
30 to 35	100	186	80	114	86	106	-20	14
35 to 40	100	186	85	116	86	102	-15	16
40 to 45	100	188	86	118	88	102	-14	18
45 to 50	100	187	89	118	87	98	-11	18
50 to 55	100	189	90	120	89	99	-10	20
55 to 60	100	191	91	123	91	101	-9	23
60 to 65	100	186	93	124	86	93	-7	24
65 to 70	100	189	96	125	89	93	-4	25
70 to 75	100	189	98	125	89	90	-2	25
75 to 80	100	186	98	125	86	88	-2	25
80 to 85	100	189	99	124	89	89	-1	24
85 to 90	100	184	98	124	84	86	-2	24
90 to 95	100	183	99	122	83	85	-1	22
95 to 100	100	182	109	132	82	74	9	32

[a] Pre-boom, bubble, and post-crash prices were measured based on annual sales price index values over the respective years 2003, 2006, and 2009. For each size category the year 2003 index value is normalized to 100.

Table 4a: Construction and Distressed Sales by Year[a]

Year	Units Built	Build/Stock	Number Distressed Sales	Distressed/All-Sales
2000	1,569	0.0562	15	0.0087
2001	1,646	0.0568	21	0.0096
2002	1,710	0.0552	35	0.0124
2003	1,766	0.0551	51	0.0149
2004	2,270	0.0678	52	0.0116
2005	2,014	0.0587	38	0.0072
2006	1,699	0.0479	18	0.0059
2007	1,084	0.0291	96	0.0438
2008	659	0.0169	879	0.3429
2009	326	0.0083	1,886	0.4638
2010	310	0.0077	1,583	0.4150
2011	285	0.0072	1,340	0.3158
2012	380	0.0095	1,045	0.2731

[a] Number of units built was calculated by differencing the total stock of housing across adjacent years, where housing stock was measured using assessment data obtained through Arizona State University. Distressed sales and all sales were measured using data obtained from the ARMLS.

Table 4b: Construction and Distressed Sales by Home Size Segment[a]

	Pre-Crash: 2000 to 2006			Post-Crash: 2007 to 2012		
Home Size Percentile	Number Units Built	Build/Stock	Number Distressed Sales	Number Units Built	Build/Stock	Number Distressed Sales
0 to 5	1,751	0.0111	163	204	0.0012	3,204
5 to 10	1,765	0.0266	68	320	0.0044	1,744
10 to 15	1,544	0.0261	50	379	0.0058	1,584
15 to 20	1,778	0.0367	40	391	0.0070	1,355
20 to 25	1,681	0.0350	41	393	0.0072	1,348
25 to 30	1,738	0.0383	32	478	0.0091	1,249
30 to 35	1,794	0.0376	36	581	0.0105	1,220
35 to 40	1,872	0.0508	26	478	0.0107	1,032
40 to 45	1,739	0.0415	26	476	0.0097	1,103
45 to 50	1,909	0.0477	24	612	0.0127	1,075
50 to 55	1,861	0.0651	19	470	0.0129	901
55 to 60	1,783	0.0622	20	498	0.0137	950
60 to 65	1,778	0.0539	20	545	0.0134	946
65 to 70	1,941	0.0656	17	527	0.0139	859
70 to 75	1,921	0.0675	15	560	0.0152	835
75 to 80	2,028	0.0815	13	606	0.0179	746
80 to 85	1,861	0.0833	11	597	0.0195	750
85 to 90	1,919	0.0913	14	615	0.0207	701
90 to 95	1,864	0.1077	12	620	0.0238	661
95 to 100	1,683	0.1069	11	802	0.0327	506

[a] Number of units built was calculated by differencing the total stock of housing for each home-size segment across adjacent years, where housing stock was measured using assessment data obtained through Arizona State University. Distressed sales and all sales were measured using data obtained from the ARMLS.

Table 5
Relative Prices Across Home Size Segments
For 5-percentile Increment Home-Size Segments From 2001 to 2012
(Dependent Variable $P_{t,s}/P_{t,50\text{-}55\ Pctl}$)[a,b]

	2001-2006	2007-2012	2001-2006	2007-2012	2007-2012
Log number of newly built homes	0.0448	0.2620	-0.0307	-0.0114	-0.0797
	(2.92)	(6.67)	(-0.80)	(-0.24)	(-0.59)
Listings/Stock in year-*t*/year-2003 (X1)	-	-	-0.5197	-2.457	-2.134
	-	-	(-1.48)	(-6.77)	(-4.57)
Log number of newly built homes * X1	-	-	0.0754	0.2753	0.2467
			(1.50)	(4.83)	(3.24)
Distressed Sales/Total Sales (X2)	-	-	-	-	-2.527
	-	-	-	-	(-1.41)
Log number of newly built homes * X2	-	-	-	-	0.2478
	-	-	-	-	(0.97)
Observations	114	114	114	114	114
Year Fixed Effects	6	6	6	6	6
R-sq within	0.173	0.544	0.194	0.716	0.741
Prob > F joint test of controls	0.033	0.0011	0.265	0.0003	0.0000

[a] Dependent variable equals $P_{t,s}/P_{t,50\text{-}55\ Pctl}$ where $P_{t,s}$ is the price index for segment *s* in year *t* and $P_{t,50\text{-}55\ Pctl}$ is the price index for the 50-55th percentile size segment (also in year *t*). Only size segments below and above the 50-55th percentile category are included in the sample. *t*-ratios in parentheses based on standard errors clustered at the year-level.
[b] ARMLS data were used to measure the number of distressed sales and homes listed. The stock of single family homes was determined using Maricopa assessment authority data. Newly built homes were calculated based on the change in stock within a given home size category between adjacent years.

Figure 1a: Repeat Sale Price Index Using ARMLS 2001:1 – 2013:9

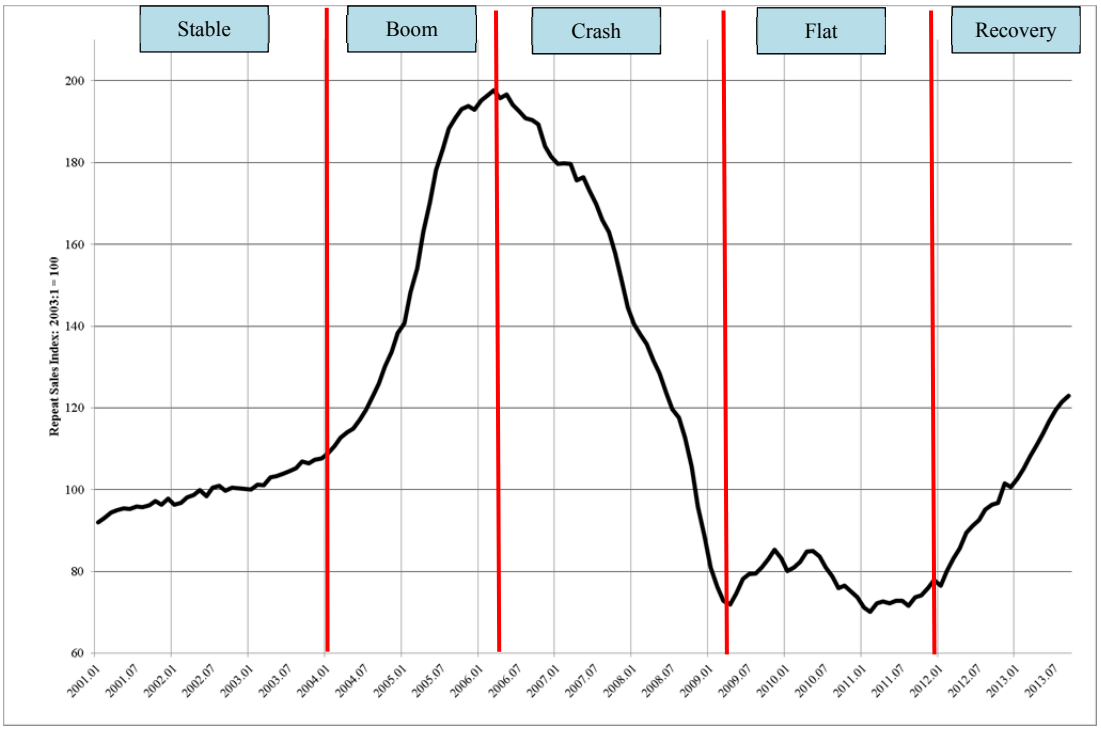


Figure 1b: Number of Homes Sold Using ARMLS 2001:1 – 2013:9

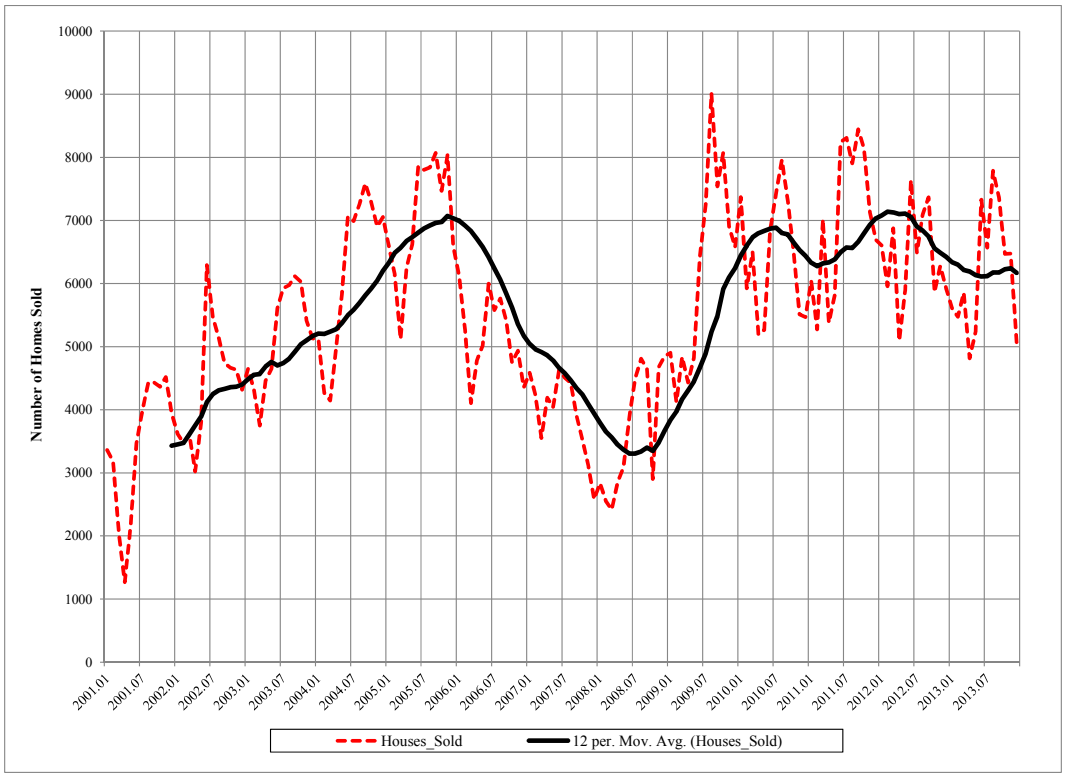


Figure 2a: Repeat Sales Indexes Using ARMLS by Home Size Segment 2001:1 – 2013:9

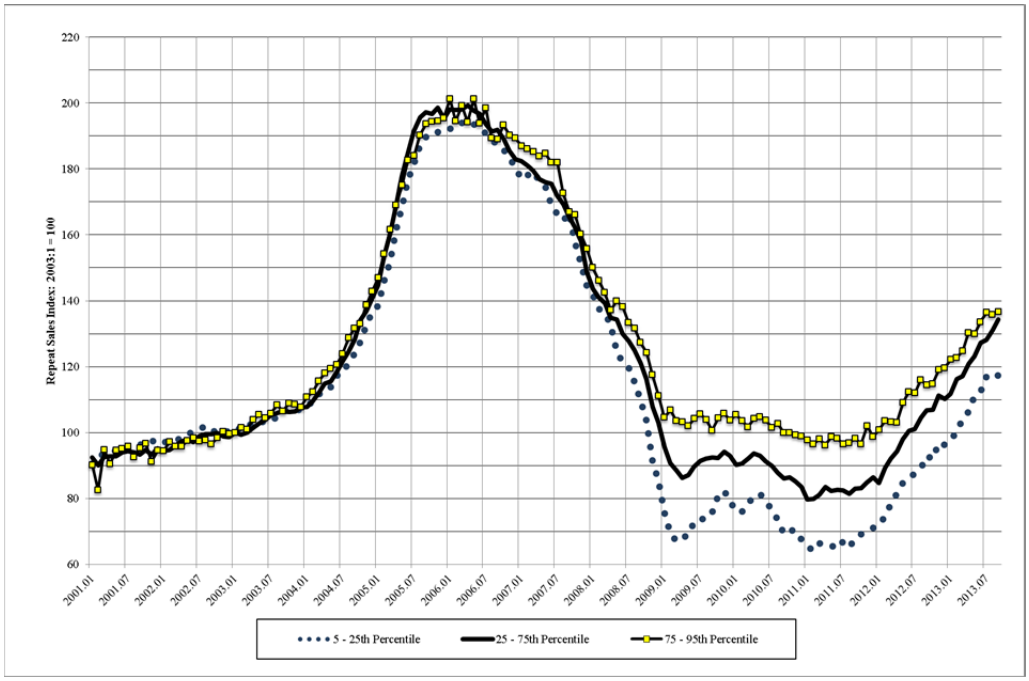


Figure 2b: Single Family Housing Starts and Repeat Sale Index 1989:1 to 2013:12
Based on FRED Data at the Saint Louis Federal Reserve Bank

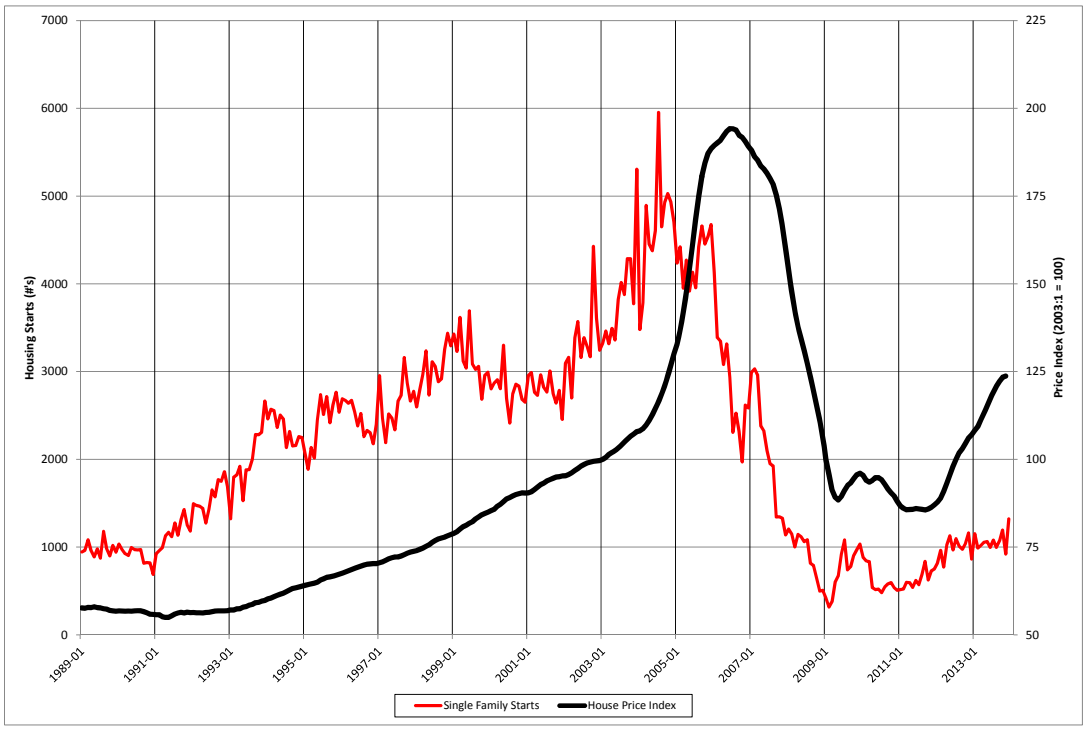


Figure 3: Repeat Sales and Repeat Rent Indexes Using ARMLS Data 2001:1 – 2013:9

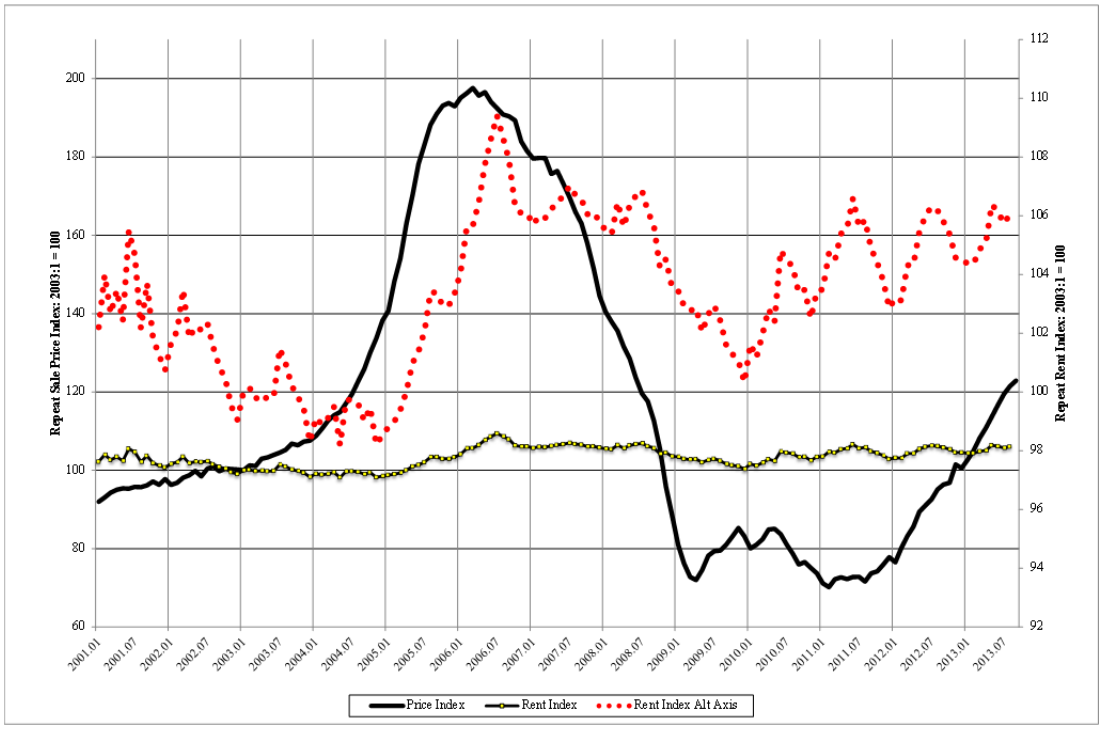


Figure 4: Repeat Sale-to-List and Sale Price Indexes Using ARMLS Data 2001:1 – 2013:9

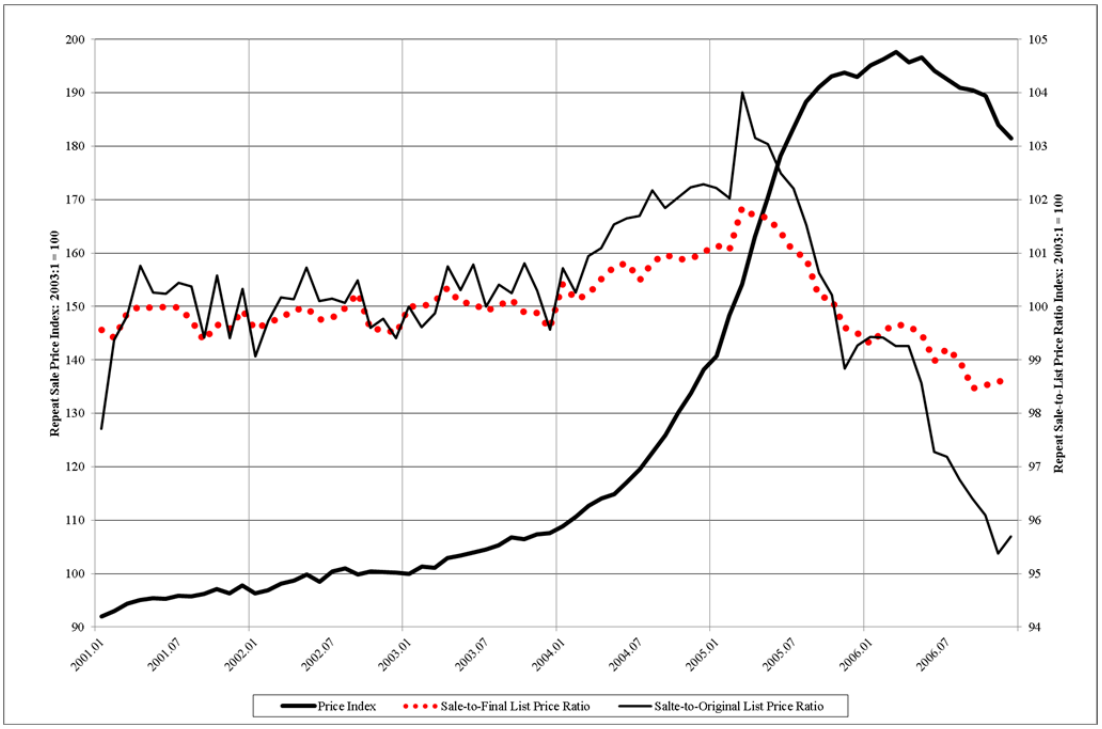


Figure 5: Median Sale-to-List Price Using ARMLS Data 2001 - 2013 by Home Size Segment (Based on Final List Price)

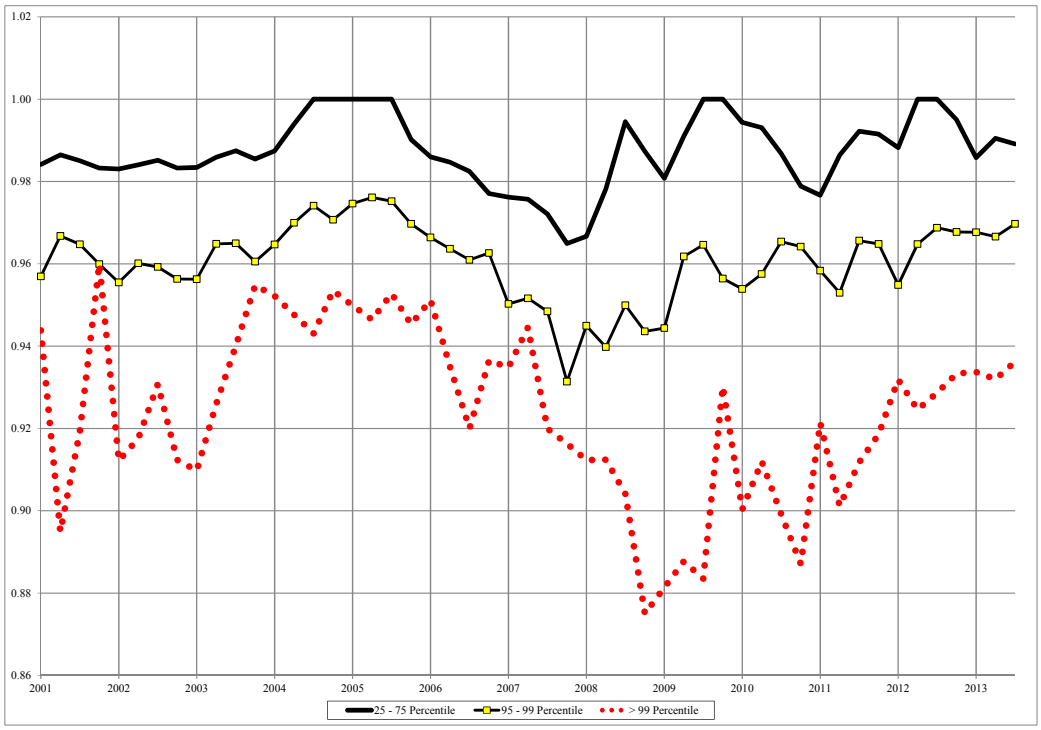


Figure 6a: Repeat Sale Price Index (Annual) Using ARMLS Data 2001-2013 Mansions Versus the Core of the Market

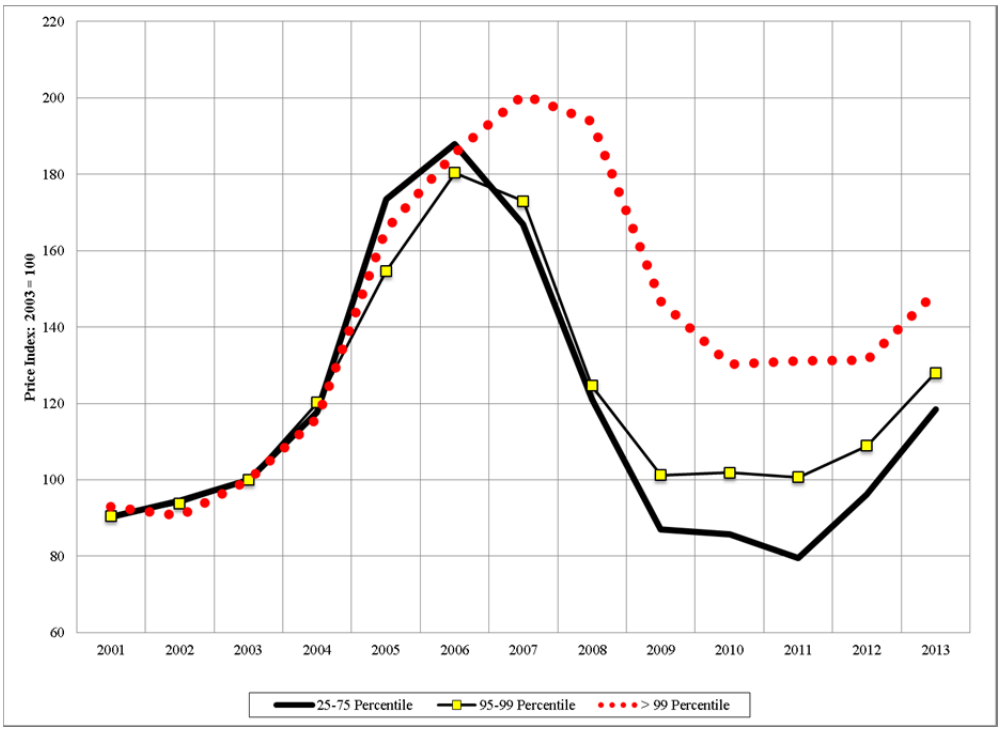


Figure 6b: Repeat Sale-to-Final List Price Ratio (Annual) Using ARMLS Data 2003-2006 Mansions Versus the Core of the Market

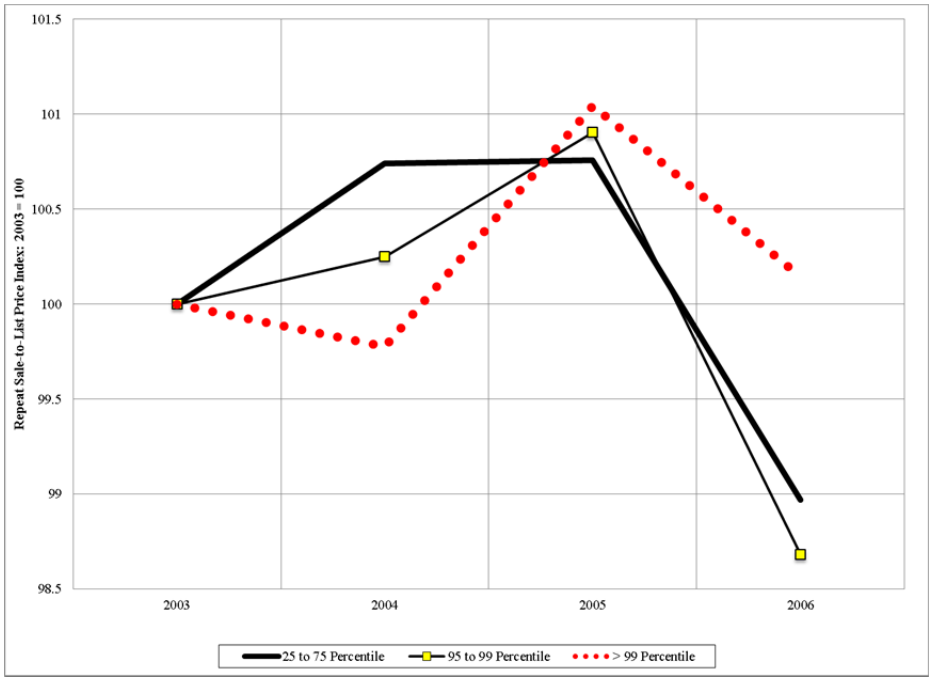


Figure 7: Post-Crash Dynamics

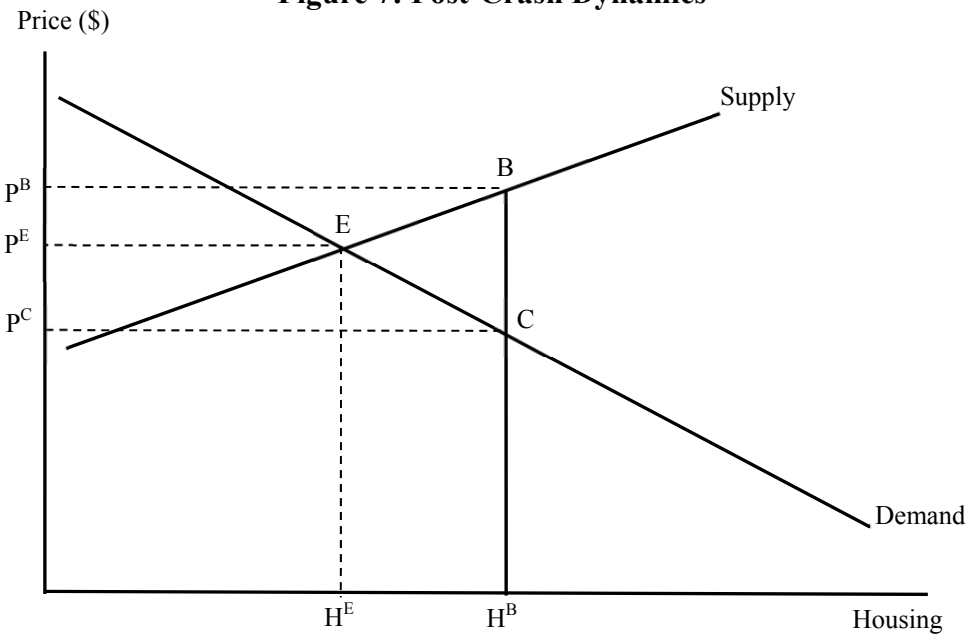
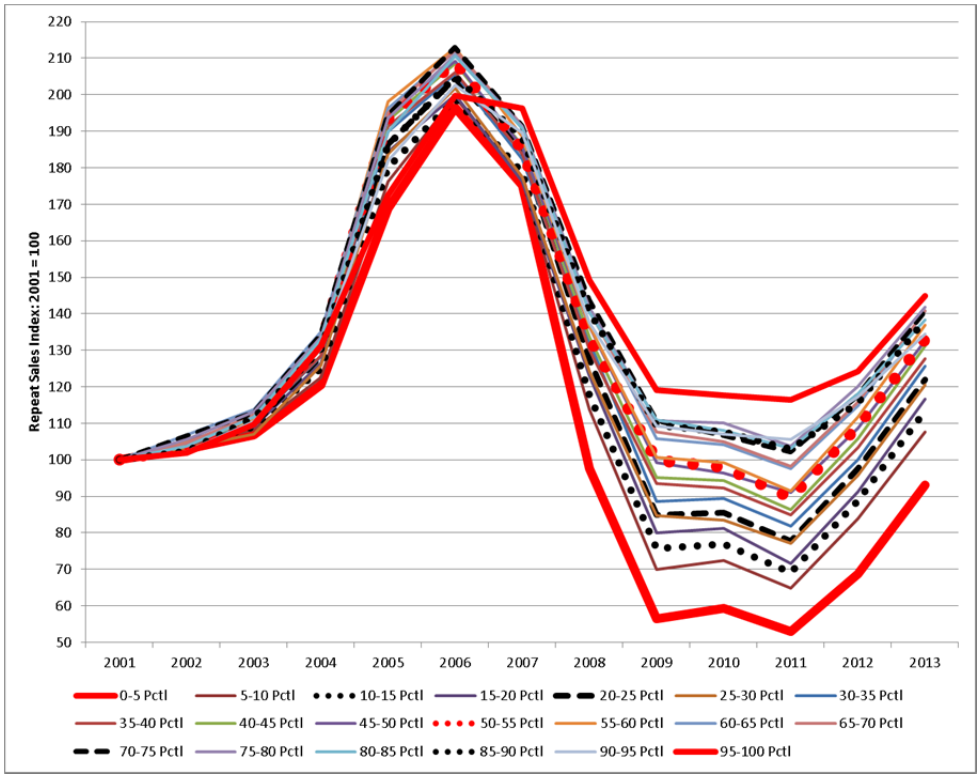


Figure 8: Annual House Price Indexes Using ARMLS Data for 20 Home-Size Segments



Appendix A: Supplemental Tables and Figures

Table A-1
Correlation Coefficients Between ARMLS and Assessment Authority Repeat Sale Indexes January 2001 to November 2011

Home Size Percentile	Monthly Indexes	Annual Indexes	Home Size Percentile	Monthly Indexes	Annual Indexes
0 to 5	0.989	0.993	50 to 55	0.986	0.997
5 to 10	0.987	0.995	55 to 60	0.983	0.995
10 to 15	0.989	0.995	60 to 65	0.987	0.991
15 to 20	0.980	0.995	65 to 70	0.973	0.996
20 to 25	0.986	0.995	70 to 75	0.978	0.990
25 to 30	0.984	0.996	75 to 80	0.965	0.985
30 to 35	0.985	0.993	80 to 85	0.961	0.994
35 to 40	0.985	0.994	85 to 90	0.983	0.990
40 to 45	0.989	0.996	90 to 95	0.967	0.984
45 to 50	0.988	0.997	95 to 100	0.945	0.981

[a] Values are the correlation between two alternate sets of home price indexes. The first set of indexes was calculated using ARMLS data which does not include FISBO ("for sale by owner") sales. The second set was calculated using Maricopa County assessment authority data and was obtained from ION corporation. The ION data reports home sales from January, 2001 through November, 2011. Although the ARMLS data was available through November 2013, for the calculations reported in this table we restricted the ARMLS sample to the same period in order to match the ION data. Both data sets include only single family homes sold.

Figure A-1: Monthly Repeat Sale Price Index Using ARMLS 2001:1–2013:9 with 95 Percent Confidence Bands

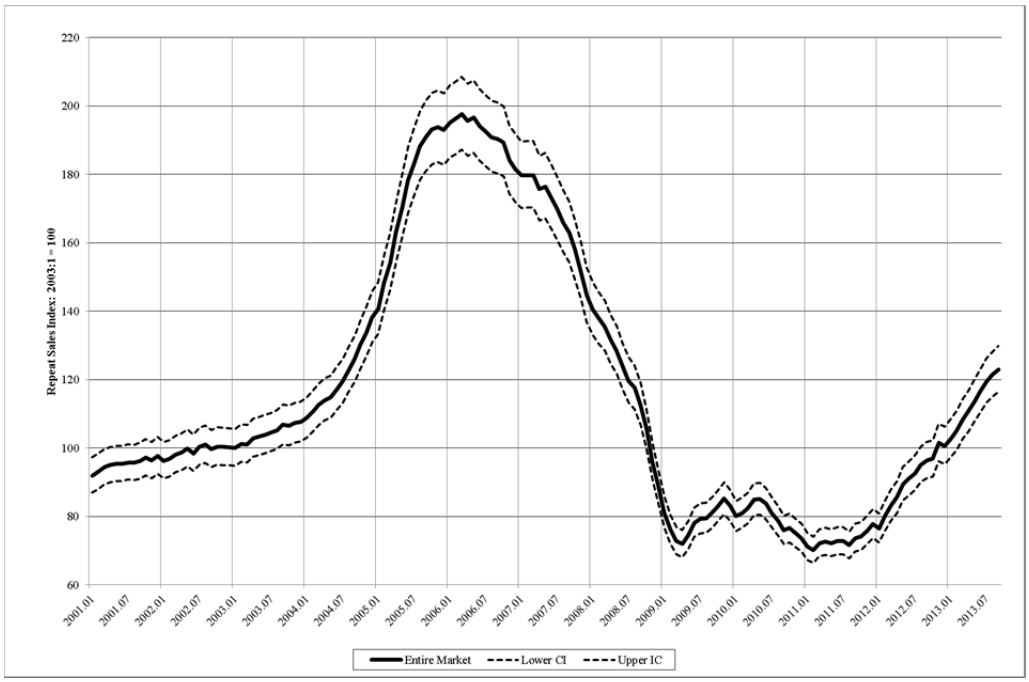


Figure A-2: Monthly Repeat Sale Price Index Using ARMLS 2001:1–2013:9 by Home-Size Segment with 95 Percent Confidence Bands

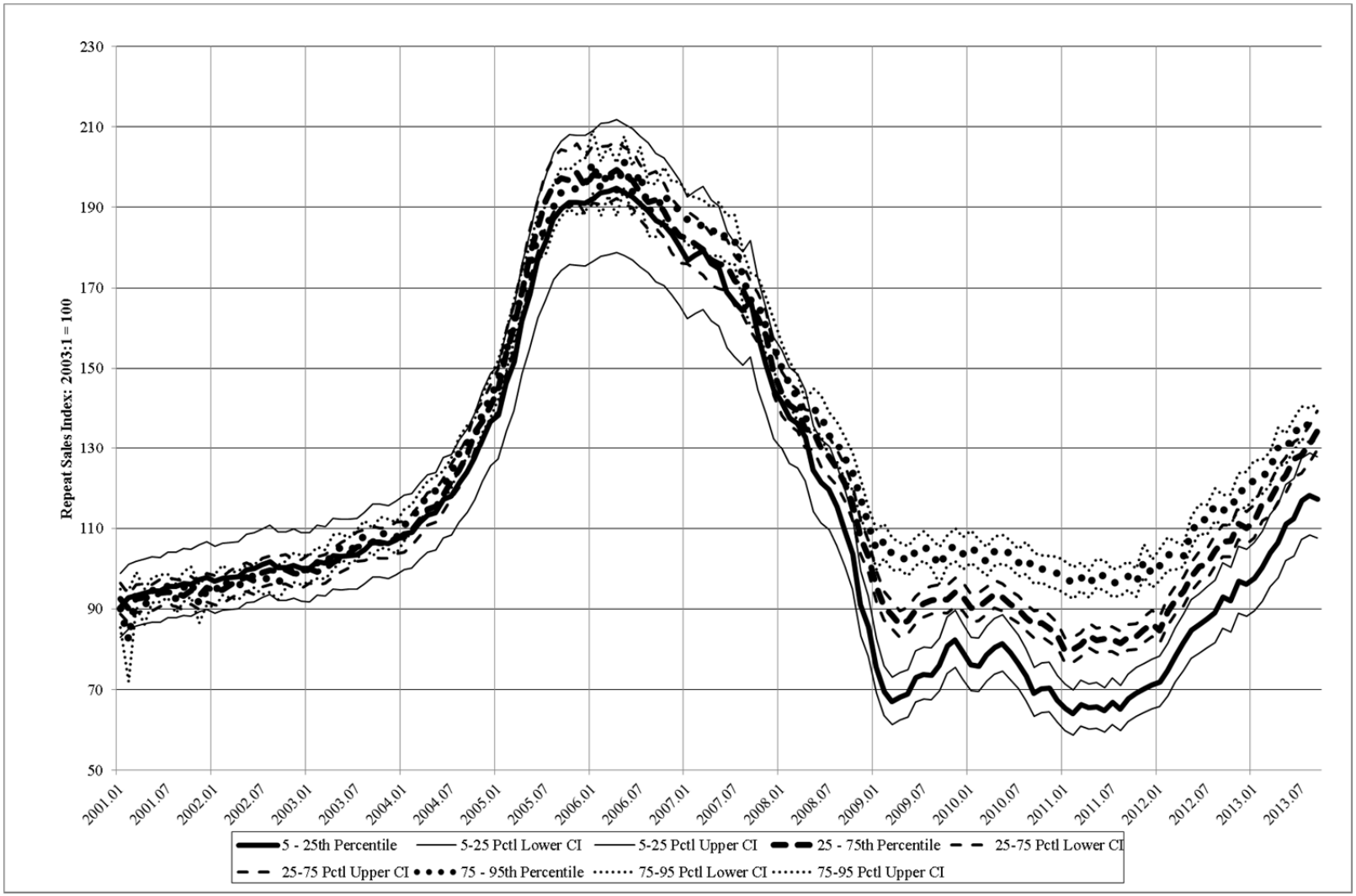
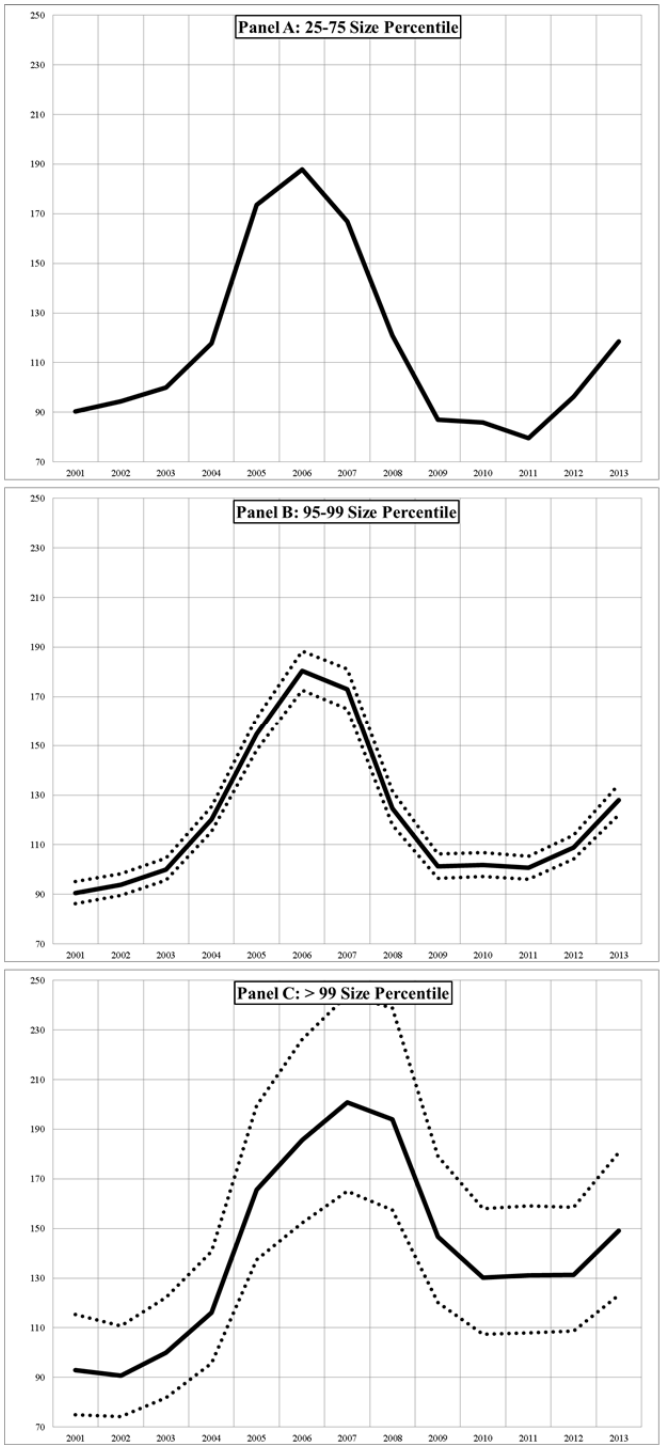
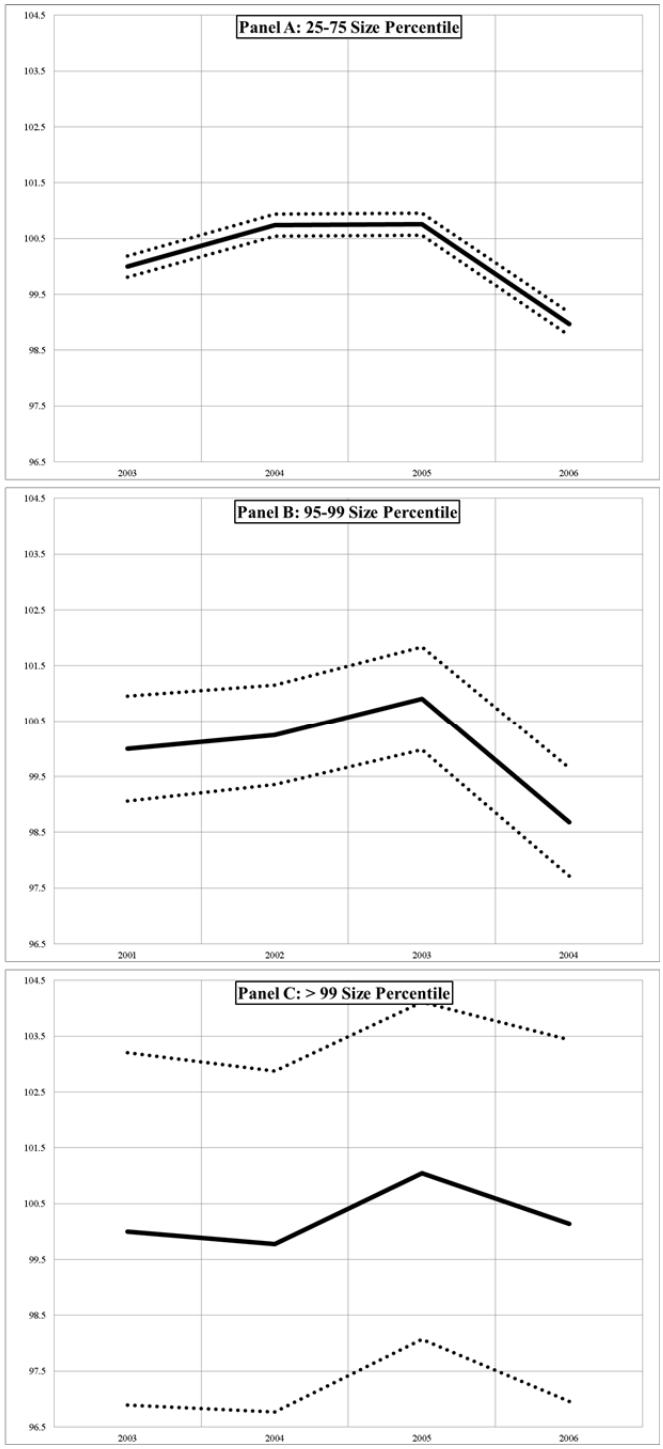


Figure A-3: Annual Repeat Sale Price and Sale-to-List Price Indexes Using ARMLS with 95 Percent Confidence Bands

Repeat Sale Price Indexes 2001-2013 (2003 = 100)[a]



Sale-to-List Price Indexes 2003-2006 (2003 = 100)



[a]Confidence bands for Panel A of the repeat sale price indexes quite narrow and do not display for that reason.

Appendix B
Access to the Data

The current study uses data from two primary sources. Our agreements with the data providers do not allow us to post or otherwise distribute the data. However, all of our data can potentially be obtained by others interested in replicating our results. Details on who to contact to obtain the data are provided below.

Arizona Regional Multiple Listing Service (ARMLS)

The ARMLS is our primary data source. The website to access the ARMLS data is http://www.armls.com/. The data are proprietary and are the property of the ARMLS, an Arizona corporation. Members of the ARMLS are the only ones who are allowed access to the data. Although we received permission to use the data in our study from the Arizona Board of Realtors, we had to ask a member to pull the data for us as ARMLS did not give us rights to access the data directly. The ARMLS is located at 130 S. Priest Drive, Suite 101, Tempe, AZ 85281-2593. Their phone number is (480)921-7777.

Maricopa County Assessor's Office Property Records

We also used data on property records from the Maricopa County Assessor's office. Two versions of the assessment authority data were obtained. The first was purchased from ION Data Express (http://www.iondataexpress.com/), a company that repackages and sells assessment data. Their email address is iondataexpress@cox.net and the phone number of their research department is (480)831-6677 extension number 15. They offer an academic price for their data. We also obtained data directly from the Maricopa County Assessor's office. Researchers who are interested in securing the assessment data should contact Linda Shaffer who is the Data Sales

Coordinator at the Maricopa County tax assessor's office. Her email is shafferl@mail.maricopa.gov and her phone number is (602) 506-7885. You will need to provide written documentation as to the nature of the research project and your school affiliation. Ms. Shaffer will either provide the data to you directly or send an email to Mary Whelan at Arizona State University who will then email you the data in a .dbf format.