Federal Housing Finance Agency Working Papers

Working Paper 12-1

Home Price Indexes for Homes in Different Price Tiers: Biases and Corrections

Andrew Leventis, Principal Economist (andrew.leventis@fhfa.gov)

Under the repeat-transactions framework for constructing house price indexes, the paper analyzes the technical challenges associated with producing unbiased price indexes for homes in distinct price tiers. The basic problem is that the "tier" to which a given home truly belongs is unobservable and can vary over time. Various approaches to forming tiered indexes, including the methodology used in the formation of S&P/Case-Shiller tiered indexes, are analyzed. Some are shown to be susceptible to mean-reversion-related biases (under- or over-statement of price changes) when home prices reflect "true" home values plus transactions-related noise. Empirical comparisons of price trends in California as a whole and two California metropolitan areas—San Francisco and San Diego—are shown where indexes are constructed under various tier-classification schemes.

Several members of FHFA staff reviewed prior drafts of this paper and provided useful comments. In particular, the author would like to thank Patrick Lawler, Austin Kelly, and Andrew Davenport for their valuable suggestions.

Background

Home price trends tend to differ along many dimensions. One important dimension is price segmentation. A number of demand and supply-side considerations, including changes in credit availability, household incomes, and other factors, can lead to divergences in price trends for expensive homes versus modestly-priced properties. Such differences, which can be significant, are of interest to economists and market observers who track real estate prices.[1] Tier-specific measures of home prices can be used to determine, for example, whether the changes in the interest rate spread between jumbo and non-jumbo mortgages differentially impact homes in different price tiers. Tier-specific price indexes can also be used to more accurately estimate the ratio of a mortgage's loan balance to the value of the collateral property. This loan-to-value ratio is a key input in statistical models that aim to predict mortgage default and prepayment.

Unfortunately, measuring price trends for homes in different price tiers is not straightforward. The fundamental problem is that a given home's "value"—and thus the price tier it belongs in—is not readily observable and, indeed, might not be thought of as a specific dollar value. When a given home is sold, a range of selling prices is possible, depending on factors such as relative buyer and seller motivation, time on market, and the number of similar properties available in the marketplace at the time of sale. In short, one could think of a home's selling price as a reflection of a home's true value plus a number of other factors, which might be considered noise.

Ultimately, for constructing home price indexes for homes in different value ranges, it is necessary to classify properties based on something observable. If home attribute information such as property square footage were available, then the classification might be performed based on attributes. In a three-tier structure, for example, homes with 0 to 1,500 square feet might be considered "low tier," those with 1,500-3,000 square feet might be considered "middle tier," and those with more than 3,000 square feet might be thought of as "high tier." Property attribute information is not available in all parts of the country, however. Also, attribute-based tiering might neglect to account for important drivers of home values. A square-footage-based approach, for example, might not perform particularly well in locations

[1] Recent descriptive literature—articles that discuss differences in trends—include: Shenk, Michael, "Metro-Area Differences in House Price Indexes," Federal Reserve Bank of Cleveland Economic Trends Report, December 11, 2008 (available at http://www.clevelandfed.org/research/trends/2008/1208/04ecoact.cfm); Maitland, Maureen, "A Look at S&P/Case-Shiller Tiered-Price Indices," S&P Dow Jones Indices Housing Views, June 23, 2011 (available at http://www.housingviews.com/2011/06/23/a-look-at-spcase-shiller-tiered-price-indices/); Humphries, Stan, "In Many Metros, Impact of Housing Downturn Depends on Value of Home," Zillow Blog, December 14, 2009 (available at http://www.zillow.com/blog/2009-12-14/in-many-metros-impact-of-housing-downturn-depends-on-value-of-home/).

where lot size is valued highly. Finally, as market fundamentals and preferences change over time, a square-footage-based bucketing scheme might become obsolete. The square footage ranges might need to change, for example, if there is a long-term drift in average home size. A 1,400 square foot house might have been a middle-tier property in the 1960s, for instance, but might very well be a low-tier home in more recent periods.

An obvious alternative to using attribute information is to use observed sales prices to "bucket" properties into value tiers. This is easy to do and avoids the data availability problems associated with the attribute-based approach. Some problems arise in the process, however, particularly when the price-based approach is used in the context of constructing repeat-transactions home price indexes. Given that the most widely-referenced set of tiered home price indexes—those released by S&P/Case-Shiller [2] —employ the repeat-transaction framework, the problems are of significant practical import.

This paper illustrates the problems that can arise in such an environment, particularly when homes are put into tiers based on a single observed transaction price. The analysis then proceeds by showing the impact of using different bucketing rules on empirical estimates of tier-specific home price indexes. Using historical sales price data for California and two cities within that state, price-tier-specific home price indexes are formed using different classifications rules. Boom- and bust-period price changes for the various approaches are then compared.

Repeat-Transactions Framework—Context

While the mechanics of the repeat-transactions methodology are beyond the scope of this paper,[3] the fundamental idea is that marketwide price trends can be measured by evaluating price changes for homes that have sold two or more times in the past. The observed price changes for specific homes, it is assumed, will provide a reasonable measure of marketwide price trends.

The basic ingredient underlying the model is the transaction pair. With the sale price information for individual properties, one can form pairs showing appreciation episodes over specific time periods. For example, if a property at 123 Elm Street sold in the first quarter of 2000 for $100,000 and then again in the fourth quarter of 2005 for $200,000, then one transaction pair is formed showing the 100 percent price growth over the 2000Q1-2005Q4

[2] These indexes are frequently cited in the media and in academic papers. See, for example, Shenk, Michael, "Metro-Area Differences in House Price Indexes," Federal Reserve Bank of Cleveland Economic Trends Report, December 11, 2008 (available at http://www.clevelandfed.org/research/trends/2008/1208/04ecoact.cfm);

[3] For a detailed description of FHFA's implementation of the repeat-transactions framework, see Calhoun, Charles, "OFHEO House Price Indexes: HPI Technical Description," available at http://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/HPI-Technical-Description.aspx.

interval. If the property then sold again for $150,000 in the third quarter of 2009, then a new pair is then formed showing a 25 percent decline over the 2005Q4 – 2009Q3 interval. These two observations, as well as other pairs for other homes, would form the dataset used for statistical analysis.

As indicated, because each of these transaction pairs shows the price change for a given property at two points in time, it is assumed that the measured appreciation (or depreciation) is a relatively unbiased measure of marketwide price trends over the identified interval. Given that millions of pairs can be formed (using data for millions of properties), the model presumes that the average price changes across all of those pairs ultimately provides reasonable measurement for marketwide price changes.

Mechanically, the repeat transactions methodology is implemented as a simple regression model where the observations are the transaction pairs. The model's dependent variable is the change in the log property price between the time of the first and second transactions in the pair. The independent variables are special dummy variables corresponding to every period in the historical time frame. For example, if quarterly data are available for 20 years, then eighty indicator variables are used as independent variables. For a specific transaction pair, all but two of the indicator variables are assigned a value of zero. The two nonzero values are those corresponding to the quarters in which the two transactions occurred. If, for example, a given transaction pair reflects sales prices from the fourth quarter of 2005 and the third quarter of 2009, then the indicator variables corresponding to those two periods would be given nonzero values[4] and all other variables would be assigned a value of zero.

The regression, which is implemented as a weighted least squares model, thus attempts to explain observed price changes using only information about when transactions occurred.[5] When estimated using a large dataset, the coefficients on the indicator variables ultimately

---

[4] The "indicator" variables are not standard 0-1 dummy variables. Rather, they take on values of -1, 0, or 1. For a given transaction pair, the variable corresponding to the date of the first transaction is given a value of -1, while the variable corresponding to the period of the second transaction is assigned a value of 1. Variables for all other periods are assigned a value of zero. Fore details, see Calhoun, Charles A., "OFHEO House Price Indexes: HPI Technical Description," available at http://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/HPI-Technical-Description.aspx.

[5] Weighted least squares is used because there is a distinct relationship between the length of time between the observed prices and modeling error; the (absolute) size of the modeling error tends to be much larger for transaction pairs with long intervals between transactions than those with short intervals. In such an environment, the precision of the coefficient (index) estimates can be improved by giving more weight to pairs with little elapsed time between transactions.

reflect the price changes that generally are observed across specific time periods. The estimated coefficients are, in effect, an index series that tracks marketwide price trends.[6]

Repeat-Transactions Framework—Illustration of the Problem

It would be easy to produce price-tier-specific home price indexes if the price tier to which each home belongs was readily observable. Unfortunately, it is not. Indeed, even if it were, homes can theoretically switch between tiers over time. If sales price are used as the basis for judging which tier a home is in, where homes transact two or more times, there will be cases where the price tier a home is in may differ across the transactions. For example, a given home's sales price may be in the bottom third of all prices in one period, but in the subsequent period, it may sell for a "middle" tier price.

While homes can certainly switch price tiers for reasons related to market fundamentals (e.g., the property neighborhood becomes more desirable over time), in many if not most cases, cross-tier switching will also be a function of ephemeral factors that influence specific transactions. An unusually "bad" outcome—where a home's sale price was particularly low because the seller was extraordinarily motivated to sell, for example—might put an otherwise middle-tier property in a low-tier price range. As discussed previously, home values are not observed and these idiosyncratic factors can be large and can easily influence the outcome (sale price) for a specific transaction.

In addressing the difficulties caused by the idiosyncratic effects, there will be implicit trade-offs. As will be shown, there are relatively straightforward ways of minimizing the effects of this noise, but those solutions come at a cost. The natural solution to the noise-related problem (and the one that is employed in this paper) amounts to bucketing properties based on the average tier they fall into, but such an approach masks information where properties switch tiers for fundamentals-related ("legitimate") reasons. For example, a property that was usually a low-tier home but that spent some time as a legitimate middle-tier property (perhaps because its neighborhood was temporarily desirable), will be branded as a low-tier home under the averaging approach. The price drift that the property experienced during its tenure as a middle-tier home thus will be incorrectly attributed to the low tier.

Whether fundamentals-based switching is large relative to noise-related switching is a subject that will be addressed in future research. In general, given the dramatic impact that idiosyncratic factors can have on observed sales prices relative to fundamental factors, it seems

[6] Because the dependent variable is the change in the *log* price of homes, to aid interpretation, the index values are formed by exponentiating the coefficients and multiplying by 100. This transformation does not affect the intuition discussed here.

likely that the noise problem is a much bigger issue. It also would seem that fundamentals-related tier switching might be mitigated by measuring price trends for small geographic areas.[7] In any case, the focus of this analysis is to illustrate the problems caused by idiosyncratic noise and to provide a solution to that problem.

A simple example can be used to illustrate the problems caused by transaction-specific noise in measuring tier-specific home price trends. The example, which is fundamentally a basic illustration of the statistical concept of "regression toward the mean," contemplates a simple world where a discrete number of homes exist. In the example, which mimics the essence of the repeat-transactions methodology, each home sells twice and tier-specific price movements are measured by calculating the average price changes for homes in each tier. Consistent with the manner in which S&P/Case-Shiller calculates its tier-specific metrics, homes are classified as low-, middle-, and high-tier based on the *first* of the two sales prices. The first sales price is compared to the distribution of all sales prices observed for the relevant period and, based on its percentile position, the home is "bucketed" as a low-, middle-, or high-tier property.

The example supposes that 999 homes exist and *true* property values are normally distributed with a mean of $200,000 and a standard deviation of $10,000.[8] In this environment, the 333 properties in each third of the value distribution are denoted as low-, middle-, and high-tier properties. While these values are "true" values, they are unobserved. Sales price reflects the true values plus noise in each of the two time periods. Noise is quantified with a zero-mean variable having a standard deviation of $10,000. The noise acts as a proxy for those idiosyncratic factors that can affect transactions prices.

Importantly, in this simple example, each home's unobserved value does not change between the first and second period; that is—there is no marketwide appreciation. For a given home, the cause of the difference in selling prices between the first and second period is thus the difference in the random noise outcome. It is assumed that the random noise component for the first time period is uncorrelated with the random noise component in the second period.

Under this scenario, despite the fact that there is zero assumed property price appreciation, it will be shown that the observed appreciation for the low-tier properties will generally be positive, the appreciation for the middle-tier properties will be near zero, and the appreciation for the high-tier properties will be observed to be negative. This results from mean reversion, which is the concept that an extreme outcome for a random variable will tend not to be

[7] The fundamentals-based switching in many cases may merely reflect changes in the relative attractiveness of certain neighborhoods (and not house-specific attributes).

[8] While the distribution of home prices generally is more similar to log-normal distribution, the normality assumption used here does not qualitatively affect the basic findings of this exercise.

followed by another extreme outcome.[9] Because of the classification system employed, in this case, there will be a correlation between the price tier a property finds itself in and the random shock that was observed. Many of the homes classified as low tier in the first period were there precisely because of an "unlucky" outcome (a negative random shock). In the second period, many of these homes will not be "unlucky" again and will be sold for middle- or even high-tier prices. For many of the low-tier properties in the first period, it thus will appear as though the properties appreciated despite the fact that the price change was merely the result of random outcomes.

The converse phenomenon will be evident at the upper end of the price spectrum. Some properties labeled as high-tier in the first period were only there by virtue of having a particularly fortuitous outcome.[10] For these homes, the noise component in the second period will not likely again be large and positive. Accordingly, because of noise alone, a share of the homes that were observed in the highest price tier in the first period will be observed as having "depreciated" in the second period.

Using the simulated 999 low-, middle-, and high-tier properties, Figures 1, 2, and 3 illustrate the phantom depreciation and appreciation problem. After random shocks are added to the properties' real values to arrive at observed selling prices in the first period, Figure 1 reveals the composition of each of the price tiers in the first period. It illustrates, for example, that among the 333 properties with the lowest selling prices in the first period, only 208 were properties that were truly low-value homes.[11] The remaining group of homes in the tier is comprised of 101 that were actually middle-tier homes and 24 that were truly high-tier homes. Notably, the high-tier properties that were observed in the lowest tier had extraordinarily poor price outcomes given their true values.

Among the properties observed with high-tier prices, the opposite (but a qualitatively similar) outcome is evident; more than 100 of the 333 properties with the highest observed sales prices were actually low and middle-tier properties that found their way into the highest tier because of idiosyncratic outcomes.

Figure 2 illustrates the transition of properties to different price tiers in the second period. The figure focuses on the transition of those properties that had low observed prices in the first period and shows their transitions to other tiers. It illustrates that among the 24 truly high-tier properties that were observed in the low-tier in the first period, only 3 remained in the low tier

---

[9] The concept of mean reversion was first discussed in Galton, Francis, "Regression Towards Mediocrity in Hereditary Stature," the Journal of the Anthropological Institute of Great Britain and Ireland, Vol. 15, pp. 246-263, 1886.

[10] For example, such an outcome would occur if a buyer overpaid for a property.

[11] As a reminder: "low-tier" properties are those whose underlying (unobserved) value is in the lowest tercile of the distribution of true home values.

in the second period. The remaining 21 properties were found to be in the middle and high tiers. Similar, but more muted transitions occurred for the 101 middle-tier properties whose sales prices were in the lowest-tier in Period 1. Among these homes, 40 transitioned to the middle tier and 25 were observed in the highest price tier in Period 2. The average observed price "appreciation" for these properties was 5.0 percent—a large number given that no actual appreciation occurred and a number that merely reflects the effects of mean reversion. Observed "appreciation" for the 24 true high-tier properties that were observed in the low tier was 8.3 percent.

Figure 3 produces "appreciation" summary statistics for each of the observed price tiers from the first period. The average price changes are shown for the 333 properties observed in the low-, middle-, and high-tiers from Period 1. As predicted, even though no price appreciation was assumed in the model, across-tier differences in average price changes are significant. The randomness in observed prices and the classification system produce price "appreciation" of 3.7 percent in the lowest price tier. For the upper tier, the opposite effect is evident—prices "fell" 3.1 percent. As would be expected, prices in the middle tier were essentially flat; in aggregate, mean reversion has no systemic effect in that tier and, as such, the observed average price change for that tier is near zero.[12]

Empirical Estimates Based on Different Classification Schemes

The simple example conveys the crux of the classification problem. One can develop the example further and simulate the effects directly in a repeat-transactions framework using a synthetic dataset of randomized home prices. Such an exercise would begin with the basic model of prices contemplated in the repeat-transactions framework, where the log of a particular home's value can be expressed as:

$$\ln\left(P_{i,t}\right) = \beta_t + H_{i,t} + N_{i,t}$$

Where

$\boldsymbol{\beta_t}$ = Baseline Home Value

$\mathbf{H_{i,t}}$ = Random Walk

[Note that, as a random walk variable, $H_{i,t} = H_{i,t-1}$+Error]

[12] Analogous to the structure of the lowest tier, in the initial period, the middle tier properties include "true" middle-tier homes as well as low- and high-tier properties that had relatively extreme outcomes. In the second period, these "misclassified" low- and high-tier properties tend to revert to their true tiers in roughly equal numbers. The observed price "decreases" and "increases" for these homes will thus tend to offset and thus the average price change for the middle tier will be close to zero.

$$\mathbf{N_{i,t}} \quad = \quad \text{White Noise such that } E(N_{i,t})=0$$

Under various assumptions for the standard deviation of the random walk and white noise components, simulations will show the same general phenomenon as was evident in the simpler example.

While simulations could be shown for various calibrated parameter estimates, ultimately the most pertinent issue is how the bucketing system affects real index estimates. Because actual transactions data are available, this section proceeds by estimating tier-specific price indexes using real data and applying different bucketing rules. Actual index estimates determined under various classification systems can then be compared.

Consistent with the simple example, price changes are analyzed for homes in three price groups—the low, middle, and high tiers. The raw data used for index construction are comprised of sales prices for homes with Fannie Mae or Freddie Mac ("Enterprise") mortgages, as well as prices for homes with FHA-endorsed mortgages and those with alternative types of financing. These data are used to publish FHFA's "expanded-data" house price indexes. Tiered indexes are not produced for FHFA's traditional HPI measures—the "all-transactions" and "purchase-only" indexes. These measures are estimated using exclusively Enterprise mortgage data.[13] As will be discussed in the Conclusion section of this paper, tier-specific versions of these indexes would require nuanced interpretation and that nuance would obscure the central point of the paper.

Given that tier-specific differences in price trends in California have been of significant interest to industry observers, the empirical work focuses on California prices. For reasons to be discussed, price-tier-specific indexes are also produced and analyzed for the San Francisco and San Diego Metropolitan Statistical Areas.

Four different classification systems are used to "bucket" homes into the three price tiers. The first of the four follows the bucketing system used by S&P/Case-Shiller. Transactions pairs are formed and each pair is classified as a low-, middle-, or high-tier home based on the value observed at the time of the first of the two transactions. Given the biases discussed in the prior sections, the expected outcome in this case is that the low-tier properties should show relatively robust appreciation in this system.

It should be noted that this approach does not guarantee that low-tier homes will show more appreciation (or less depreciation) than other tiers. Rather, the relevant point is that the low tier will show more appreciation (or less depreciation) than it would if measured with an

[13] The "all-transactions" index is estimated using appraisal values from refinance mortgages as well as sales prices from purchase-money mortgages. The "purchase-only" measure relies exclusively on data from purchase-money mortgages.

alternative bucketing scheme. In other words, the relevant comparison is not how a given tier's index compares with other tiers, but rather how the across-tier differences vary under the alternative classification schemes.

Other than the potential for bias, another notable characteristic of the first approach is that it applies to pairs and not to homes *per se*. The approach classifies a given *pair* as being in the low, middle, or high tiers and thus only applies to a home for the two transactions referenced in the pair. To the extent that a given home may be associated with multiple transaction pairs,[14] the tier positions for different pairs are set independently. A specific pairing may be classified as high-tier even if, for the same home, five preceding and subsequent pairs give it a low-tier classification.

The second of the four classification systems is simply the opposite of the first: the bucketing is performed based on the tier position of the home at the time of the second transaction. This approach has the opposite bias as the first. The mean reversion phenomenon coupled with the bucketing timing will understate price appreciation in the low tier and overstate it in the high tier. Importantly, during the bust period, this approach will clearly overstate the price declines evident for low-tier properties; homes that are observed with the largest negative price shocks will disproportionately be represented in that tier and thus it will appear that the low-tier properties experience relatively large price declines.

The third and fourth classification approaches attempt to mitigate the mean-reversion problem. The third is based on published research from 1993. Recognizing the mean-reversion problem illustrated in this paper, an article by Chris Mayer[15] proposed a classification system that analyzed time-normalized prices. Under the proposal, the two observed prices in each transaction pair are converted to equivalent values in a specific historical time period using a local-area house price index. The two normalized values would then be averaged and compared to the distribution of all averages (across all pairs).

A simple example can illustrate the approach. Suppose that three time periods exist and that an aggregate home price index (formed using all available data) has values of 100, 110, and 120 in those three periods. Then suppose the three transaction pairs are observed as below:

[14] This will occur where properties have transacted three or more times in the past.
[15] See Mayer, Chris, "Taxes, Income Distribution, and the Real Estate Cycle: Why All Houses Do Not Appreciate at the Same Rate," New England Economic Review, May/June 1993, available at: http://www.bos.frb.org/economic/neer/neer1993/neer393c.pdf. See, in particular, footnote 17.

	Period of Transaction 1	Trans. Price 1	Period of Transaction 2	Trans. Price 2
	[a]	*[b]*	*[c]*	*[d]*
[1]	1	$50,000	2	$65,000
[2]	2	$65,000	3	$81,000
[3]	1	$40,000	3	$41,000

For each of these three pairs, the approach begins by converting the transactions prices to the equivalent Period 1 values using the price index. So, for example, the first and second transactions in the first pairing have equivalent first-period values of $50,000[16] and $59,091[17] respectively. The second pairing has equivalent first-period values of $59,091 and $67,500. The table below illustrates this example.

	Period of Transaction 1	Trans. Price 1	Period of Transaction 2	Trans. Price 2	Normalized Price Trans. 1	Normalized Price Trans. 2	Avg. of Normalized Prices
	[a]	*[b]*	*[c]*	*[d]*	*[e]*	*[f]*	*[g]*
[1]	1	$50,000	2	$65,000	$ 50,000	$ 59,091	$ 54,545
[2]	2	$65,000	3	$81,000	$ 59,091	$ 67,500	$ 63,295
[3]	1	$40,000	3	$41,000	$ 40,000	$ 34,167	$ 37,083

After the average normalized values (column [g] above) are computed for each of the pairs, the pairs are then classified in the low, middle, or high tiers based upon where its normalized average value lies with respect to the distribution of all computed average values (column [g]). Tier-specific price indexes are then separately estimated using those pairs that are in the respective tiers.

Because this approach uses both the first and second transactions prices in the classification process, much of the mean-reversion-related bias that afflicts the other bucketing schemes is mitigated. The tier-specific indexes that would be produced would be relatively "clean." Unfortunately, the approach treats each pairing independently and thus does not always make full use of property-level data. A given property may be associated with multiple pairings and, while the third approach reasonably deals with the mean-reversion problem, it is somewhat myopic.

[16] $50,000 = $50,000 * (100/100).

[17] $59,091 = $65,000 * (100/110)

The fourth classification system, which is a simple offshoot of the Mayer approach, attempts to solve the mean-reversion problem while being *property*-focused (as opposed to pair-focused). The focus on bucketing *properties* instead of *pairs* addresses the problem that arises when there is an anomalous set of transaction prices for a given home. If a given home is demonstrably a low-tier property based on a great deal of data, but yet has an anomalous pair of transaction prices, then the first three approaches will sometimes make use of misclassified data.

The third approach is less susceptible to misclassification than the other two, to be sure, but the vulnerability still exists. Even if one of the two transactions in a transaction pair reflects a reasonable value for a given property, a particularly anomalous outcome for the other price will affect the average normalized price. The basis for classification is merely the midpoint between two normalized prices and that midpoint will be influenced by anomalous observations. Thus, misclassification can occur.

The fourth approach determines a property's classification based on the average normalized price across all available prices for the property. The table below adds property address information to the prior example and illustrates the new approach.

	Address	Period of Transaction 1	Trans. Price 1	Period of Transaction 2	Trans. Price 2	Normalized Price Trans. 1	Normalized Price Trans. 2	Avg. of Normalized Prices
		[a]	*[b]*	*[c]*	*[d]*	*[e]*	*[f]*	*[g]*
[1]	123 Elm	1	$50,000	2	$65,000	$ 50,000	$ 59,091	$58,864
[2]	123 Elm	2	$65,000	3	$81,000	$ 59,091	$ 67,500	
[3]	55 First St.	1	$40,000	3	$41,000	$ 40,000	$ 34,167	$ 37,083

The first two transaction pairs in the prior example are shown to be for the same property: 123 Elm Street. The relevant average price in this case will be the average price across the three available transactions for 123 Elm: $50,000, $59,091, and $67,500. For classification purposes, that average will then be compared to the average for 55 First Street (which only has two transactions) as well as averages for other available properties.

This fourth approach forces a home to be in one price tier through all of time; price trends observed for a given home will generally be attributed to the price tier in which it was most commonly found. As outlined earlier, this approach rigorously addresses the mean reversion problem, but does so at the expense of losing information. Where properties have temporary transitions from one price tier to another for fundamentals-related reasons, the price trends evidenced during the temporary period will be ascribed to the "wrong" price tier.

Empirical Result: California

Using the four classification schemes, Figure 4 reports tier-specific price indexes for the state of California. The reported indexes, which are constructed for the low-, middle-, and high-price tier properties, are set to 100 in the first quarter of 2000. By basing each of the indexes to 100 in that year, the relative magnitude of boom-period price run-up and subsequent collapse are clearly displayed.

The first two panels in the figure, which correspond to the first-period and second-period bucketing schemes, clearly illustrate the extent to which the classification system affects estimates. For example, under the two approaches, the evolution of prices for low-tier homes differs substantially. Under the first-period bucketing system (the method used by S&P/Case-Shiller), the low-tier properties appreciate by a much greater percentage than others during the boom and then return to price levels consistent with what is observed for the other tiers. By contrast, under the second-price classification system, the boom-period appreciation for low-tier properties is about the same as for the middle tier but then values collapse. By the second quarter of 2012, the second-period bucketing approach estimates that prices for low-tier homes were generally at 2000 levels. Under the first-period bucketing, 2012Q2 prices for low-tier homes remained roughly 30 percent above 2000 levels.

The third and fourth panels in Figure 4 show tier-specific indexes formed under the third approach (henceforth the "Mayer approach") as well as the fourth method (henceforth the "modified-Mayer approach"). Perhaps not surprisingly, the Mayer approach generally produces estimates that are in between those shown in the first two panels. While the first two approaches estimated low-tier appreciation of 198 and 165 percent respectively between 2000 and the peak, the Mayer approach calculates a 181 percent increase. Similarly, the price change between the peak and the second quarter of 2012 for the low-tier properties is measured at 58 percent under the Mayer approach, between the 55 and 61 percent estimated under the other methods.

Interestingly, a comparison of the third and fourth panels in the figure indicates that the methods produce almost identical results. It appears that the problem associated with misclassification of homes (alternately stated--the benefit associated with allowing the classification to be informed by more than two sales prices) is small; the index values produced under the two systems are generally within a point or two of each other.

Table 1 summarizes the results from the S&P/Case-Shiller bucketing approach and those from the modified-Mayer (fourth) method. The results indicate that the differences between the two bucketing schemes, although not dramatic, are material. The differences are particularly notable over longer time frames, which is an intuitive outcome given that the mean-reversion bias will tend to compound over time; the two approaches' respective estimates of aggregate price change are most different when analyzed over the entire interval between the first quarter of 2000 and the second quarter of 2012.[18] Also not surprising is the fact that differences between the first-transaction bucketing and the modified-Mayer approaches are relatively muted for the middle-tier properties. While the potential impact of correcting for mean-reversion is clear for low- and high-tier homes, for middle-tier properties the directional impact is not obvious.

Empirical Results: San Francisco and San Diego Metropolitan Statistical Area

Unless tier-specific home price indexes are estimated for extremely small geographic areas, an important question will inevitably arise in connection with such metrics: "Don't tier-specific price indexes overweight price trends in neighborhoods with large representations of homes in the identified tiers?" The tier-specific indexes reflect price trends for homes in a given price tier, but also necessarily include neighborhood effects; to the extent that homes tend to be clustered in neighborhoods with similarly-priced properties, the tier-specific price indexes can become *de facto* indexes for certain sub-areas. In the California example above, for instance, the low-tier index is largely comprised of properties in the Central Valley. Thus, although the index incorporates price trends for some low-price properties in other parts of the state, it may look very similar to an index that would be produced with all properties in the Central Valley.

Unless tier-specific price indexes could be formed for individual streets—clearly an impractical option—tier-specific price indexes will necessarily reflect both neighborhood effects as well as the actual effect of being in the denoted price tier. Tiered price indexes will tend to be more reflective of the "pure" price effects, however, when smaller geographic areas are the basis for index estimation. In this section, to more closely isolate the price-tier effects, indexes are estimated at the city level. Price-tier indexes for San Francisco Metropolitan Statistical Area (MSA) are the focus of the analysis.

Mirroring the structure of the results summary shown in Table 1, Table 2 shows price-tier-specific price changes determined for San Francisco under the different classification approaches. Although the boom- and bust-period price changes for San Francisco's homes differ from those reported for California, the same qualitative difference between the classification schemes is apparent. The low-tier San Francisco properties appreciated the most during the boom, although—as with the statewide results—the measured rate of increase is diminished under the modified-Mayer bucketing approach. Also consistent with the results for

[18] The second quarter of 2012 represents the most recent available data.

California is the fact the impact of switching classification schemes is greatest for the low- and high-tier homes and relatively modest for the middle tier.

In case the findings for San Francisco represent an empirical anomaly, tier-specific indexes have been estimated for San Diego. The results for San Diego are qualitatively the same as for San Francisco. Table 3 shows that, consistent with the results for San Francisco, the bucketing approach described in this paper reduces the measured amplitude of the boom and bust for the low- and high-tier indexes.

Empirical Results: Comparisons with S&P/Case-Shiller-Tiered Indexes

The S&P/Case-Shiller tier-specific indexes are produced using a slightly different dataset from that used by FHFA.[19] Also, the version of the basic repeat-transaction model that it uses slightly differs from the version employed by FHFA.[20] Those differences notwithstanding, it is still interesting to compare the respective tier-specific indexes.

For both low and high-tier properties, Figure 5 compares the S&P/Case-Shiller metrics with those produced in this paper for San Francisco. In particular, three indexes are shown for each price tier: (1) the S&P/Case-Shiller index, (2) the FHFA index formed using the classification rule employed by S&P/Case-Shiller, and (3) the FHFA index formed with the modified-Mayer approach.[21] To facilitate comparisons, the FHFA indexes are shown in green. Also, the two indexes that rely on the S&P/Case-Shiller-like bucketing are graphed with uninterrupted lines, while the FHFA index with the new bucketing scheme is shown with a dashed line.

Although it is not possible to know with certainty what the S&P/Case-Shiller indexes would look like under a different classification system, one crude (but reasonable) assumption would be that the gap between the respective S&P/Case-Shiller metrics would resemble the gap between the presented FHFA indexes. In other words, the divergences shown for the FHFA indexes, when coupled with the existing metric from S&P/Case-Shiller, give us a clue as to what the S&P/Case-Shiller metric would look like under an alternative classification scheme.

Under the modified-Mayer approach to classification, the figure suggests that the S&P/Case-Shiller low-tier index would likely evidence more modest boom-period price appreciation than was previously measured. While the as-stated S&P/Case-Shiller low-tier index grew roughly 163 percent between 2000 and 2006, assuming the S&P/Case-Shiller adjustment resembles the

[19] The S&P/Case-Shiller indexes are formed with sales price information recorded at county recorder offices.

[20] For a discussion of the methodological and data differences, see Leventis, Andrew, "Revisiting the Differences between the OFHEO and S&P/Case-Shiller House price Indexes: New Explanations," OFHEO Research Paper, January 15, 2008, available at: http://www.fhfa.gov/PolicyProgramsResearch/Research/Pages/Revisiting-the-Differences-between-the-OFHEO-and-SPCase-Shiller-House-Price-Indexes-New-Explanations.aspx.

[21] The FHFA indexes used here are those that were presented in the prior section.

FHFA-based adjustment, an adjusted S&P/Case-Shiller index might show an increase of about 148 percent.[22]

The S&P/Case-Shiller estimates for high-tier properties would likely evidence more modest declines in the recent bust. With the FHFA data, the modified Mayer approach produced bust-period declines that were only 71 percent as large as those estimated under the first-transaction bucketing. Applying that 71 percent factor to the observed 22 percent decline in the S&P/Case-Shiller high-tier index, one might assume that the decline for an alternatively-calculated S&P/Case-Shiller high-tier index might be roughly 16 percent.

Conclusion

Although the S&P/Case-Shiller indexes are the most widely-cited tier-specific metrics, other metrics are available. Zillow, for example, provides some tier-specific metrics on its website and in periodic reports. Its metrics are produced by bucketing properties based on their latest estimated value (its latest "Zestimate"). To the extent that the latest estimates probably are heavily influenced by the latest sales price for a given property, the tier-specific estimates may be subject to the same type of bias as was observed for the second-period bucketing approach described earlier. The magnitude of the bias may be smaller, however, because the Zillow indexing procedure does not appear to focus on the repeat-transactions methodology used by FHFA and S&P/Case-Shiller.[23]

Proprietary tier-specific indexes are sold by CoreLogic, Loan Processing Systems, and Fiserv. Specific methodological details describing how these entities categorize properties appear not to be available to the general public, however. Thus it is not known whether and how these metrics are influenced by biases caused by mean reversion.

FHFA currently releases several types of indexes with various datasets and for various levels of geographic aggregation. FHFA tiered indexes, such as some of those in this paper, could also be produced. Among FHFA's "purchase-only," "all-transactions," and "expanded-data" metrics, it seems clear that the "expanded-data" price indexes would be the most obvious candidates for tiering. Those indexes have full coverage along the price spectrum and thus the low-, middle-

[22] For the FHFA indexes, the 2000Q1-Peak price increase was measured to be 156.3 percent under the first-transaction bucketing scheme and 141.4 percent under the modified-Mayer approach. The modified-Mayer approach thus produced an aggregate appreciation that was about 90 percent of the appreciation estimated under the original bucketing. The reported figure of 148 percent was computed as 90 percent (=141.4/156.3) of the 163.2 percent appreciation observed under the standard S&P/Case-Shiller bucketing system.

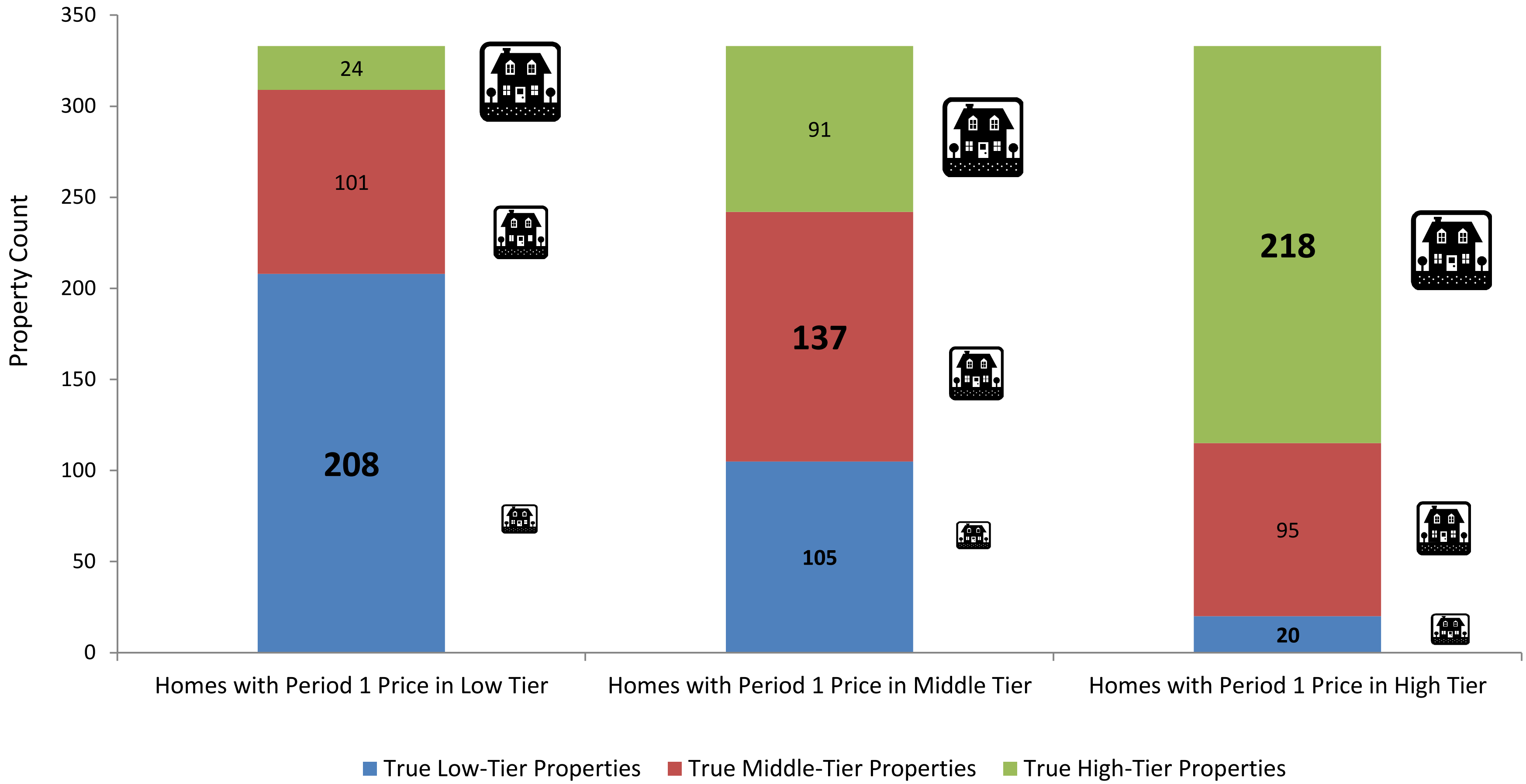
[23] Zillow's methodology documents suggest that hedonic indexing plays a key role in its statistical model. No specific details are published, however. Some information can be found, however, at http://www.zillow.com/wikipages/What-is-a-Zestimate/.

and high-tier indexes produced with such data would be most complete and meaningful.[24] Many metropolitan areas would not have enough data to facilitate production of reliable tiered metrics, however. Also, although larger geographic aggregations (e.g. states) would tend to have sufficient data, as discussed in this paper, the interpretation of tiered metrics at such high levels of aggregation is troublesome given the geographic clustering of homes in the same price tier.

[24] The interpretation of tier-specific indexes produced exclusively with the Fannie Mae and Freddie Mac data-- i.e., the all-transactions and purchase-only indexes--would be strained. The conforming loan limit makes those datasets underrepresentative of expensive homes and thus a "high-tier" index constructed with such data might not reliably reflect price trends for the most expensive homes in the market. Rather, the high-tier index would reflect price trends for the most expensive properties *financed by Fannie Mae and Freddie Mac*.

Figure 1: Distribution of Homes in First Period by "True" Value

Where Home Values are Observed *with Some Error, "Misclassification" will Occur*



* - Observed Value = Actual Value + Random Variable (Zero Mean, Std. Dev=$10,000)

Figure 2: Transition of Low-Tier Homes into Second-Period Tiers

Most Misclassified Properties Return to Their Appropriate Tiers

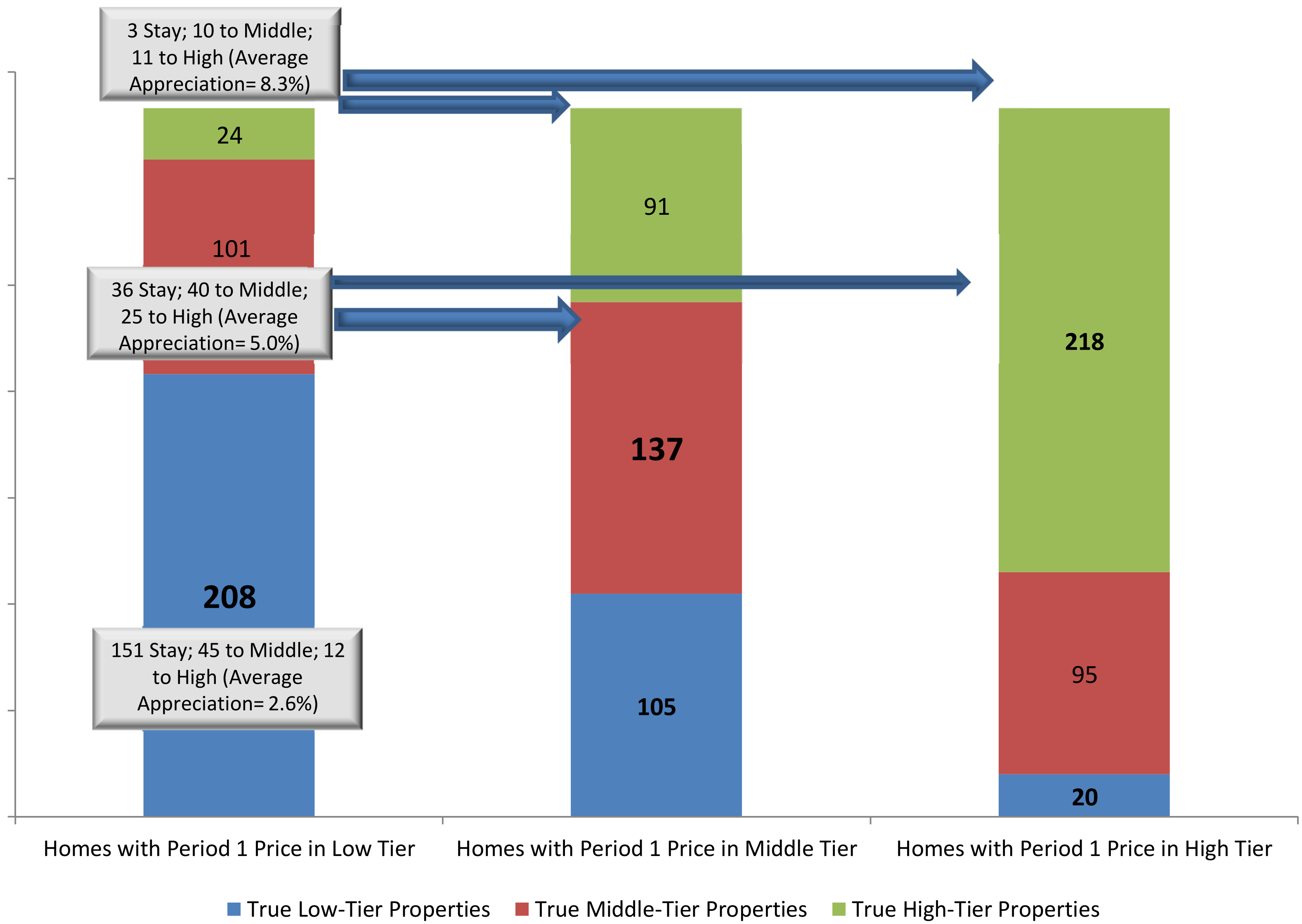


Figure 3: Average Observed Price Change for Homes in Different First-Period Price Tiers

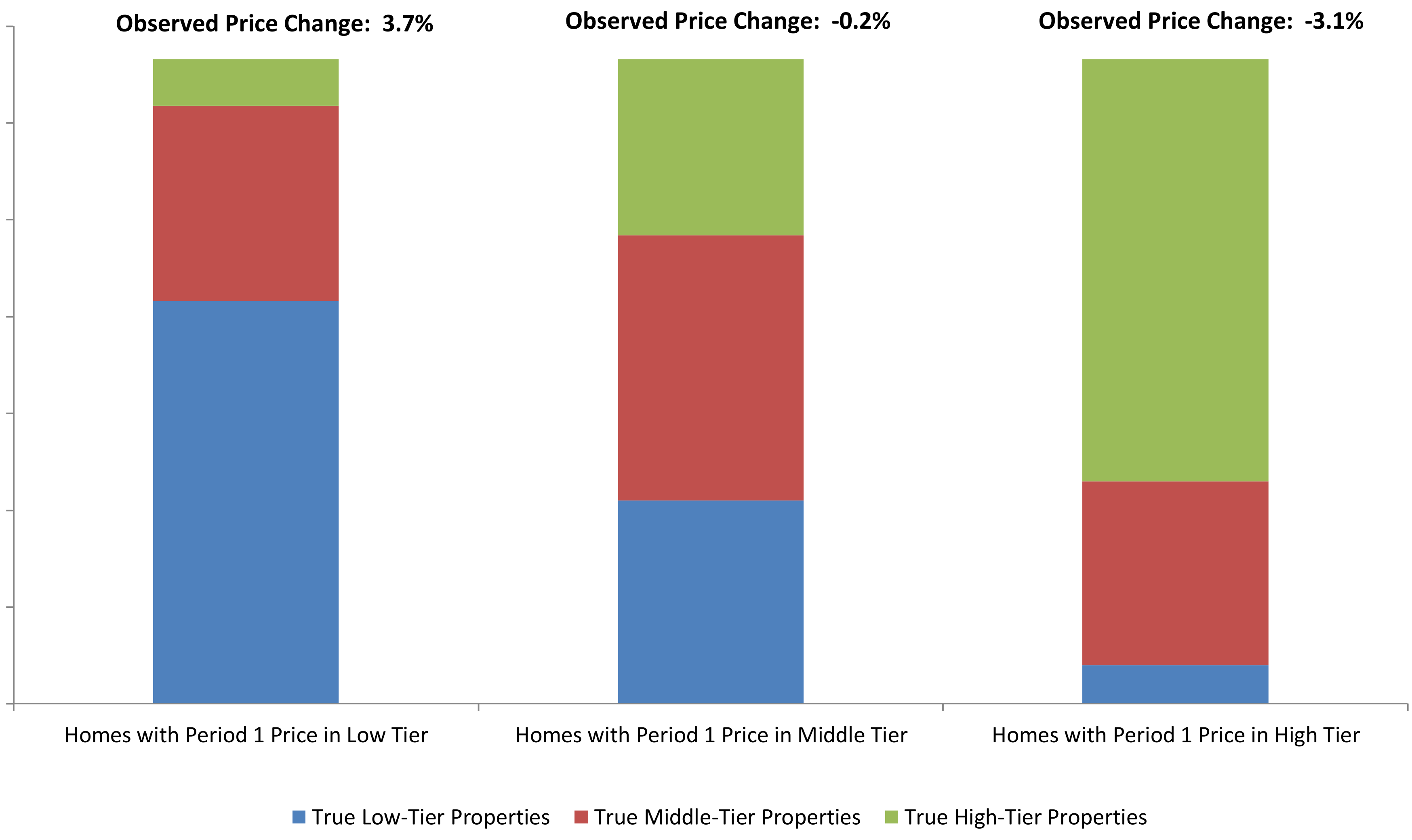


Figure 4: Tier-Specific Price Indexes Calculated for California under Various Approaches
(California; FHFA "Expanded-Data" House Price Index)

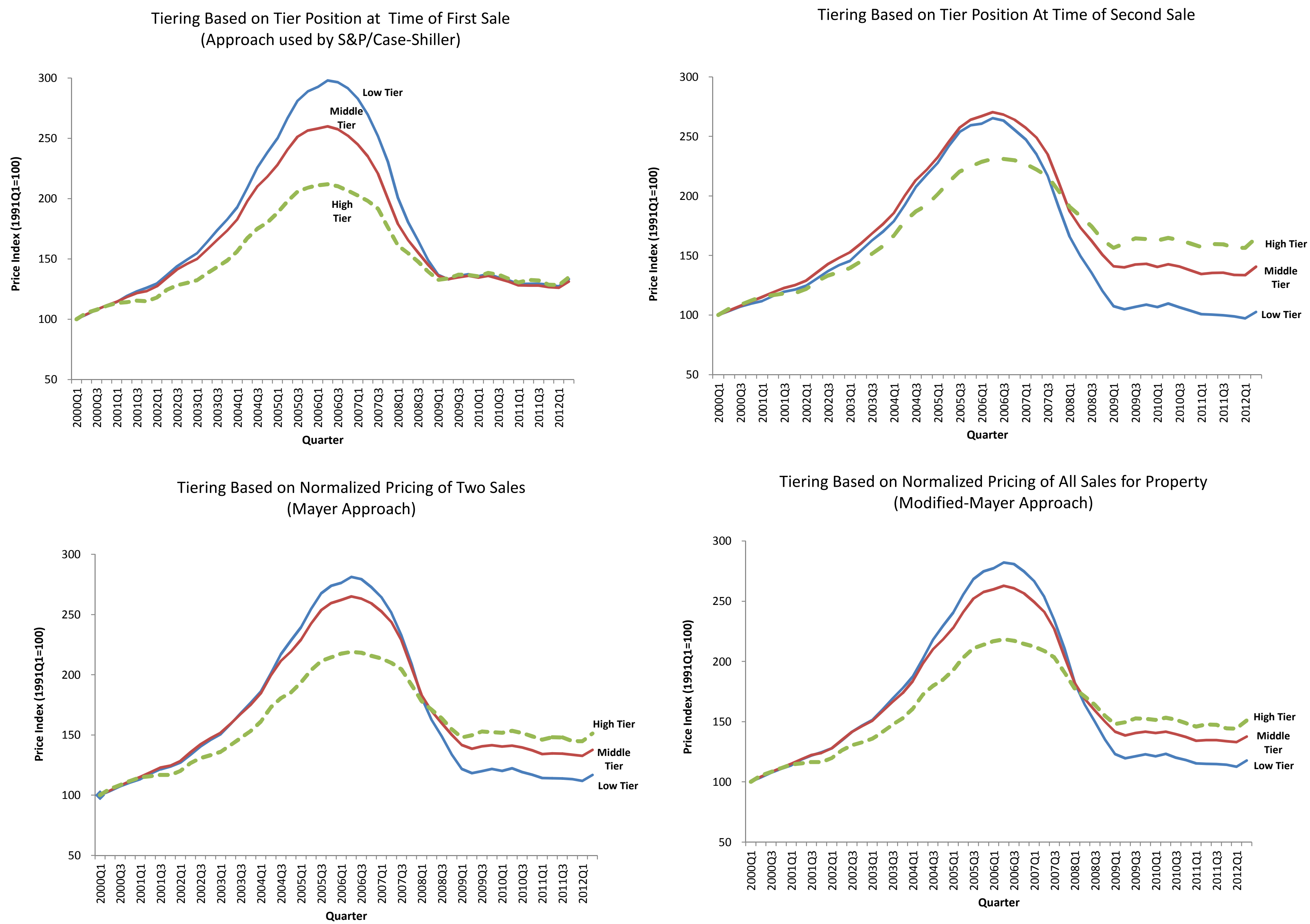
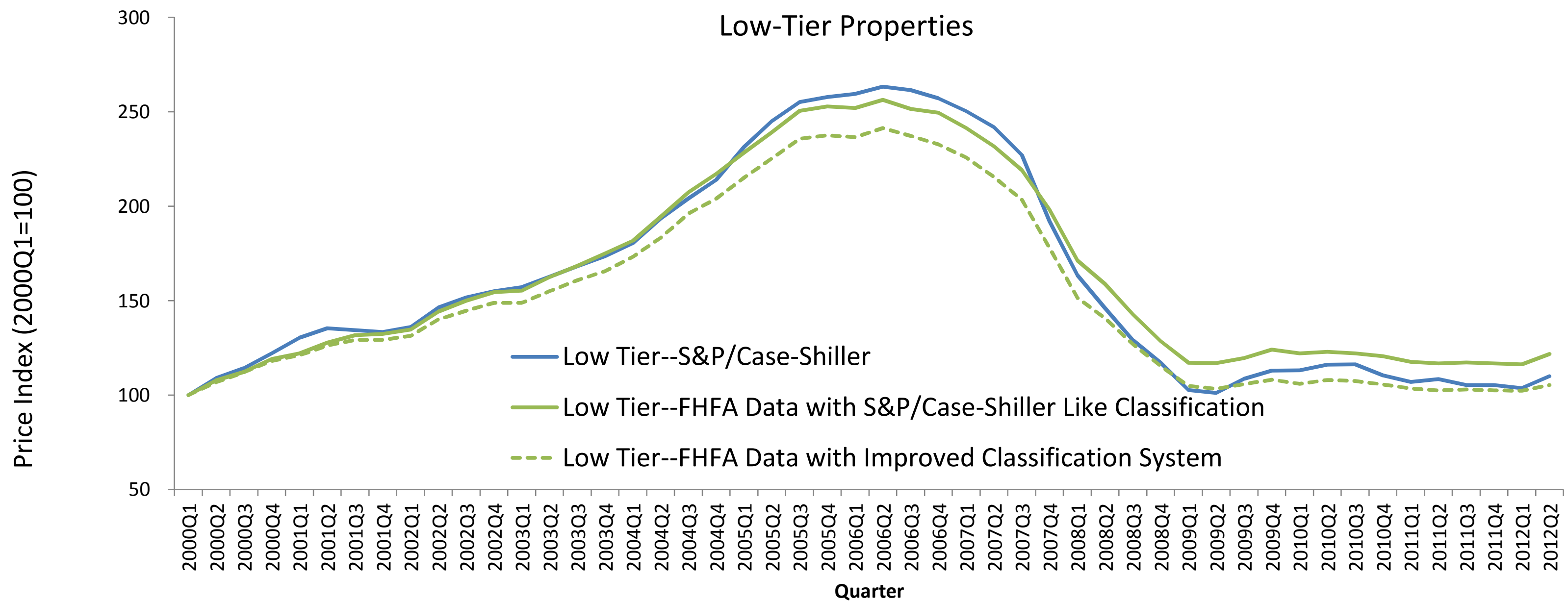


Figure 5: Tier-Specific Price Indexes Calculated for San Francisco
by S&P/Case-Shiller and FHFA



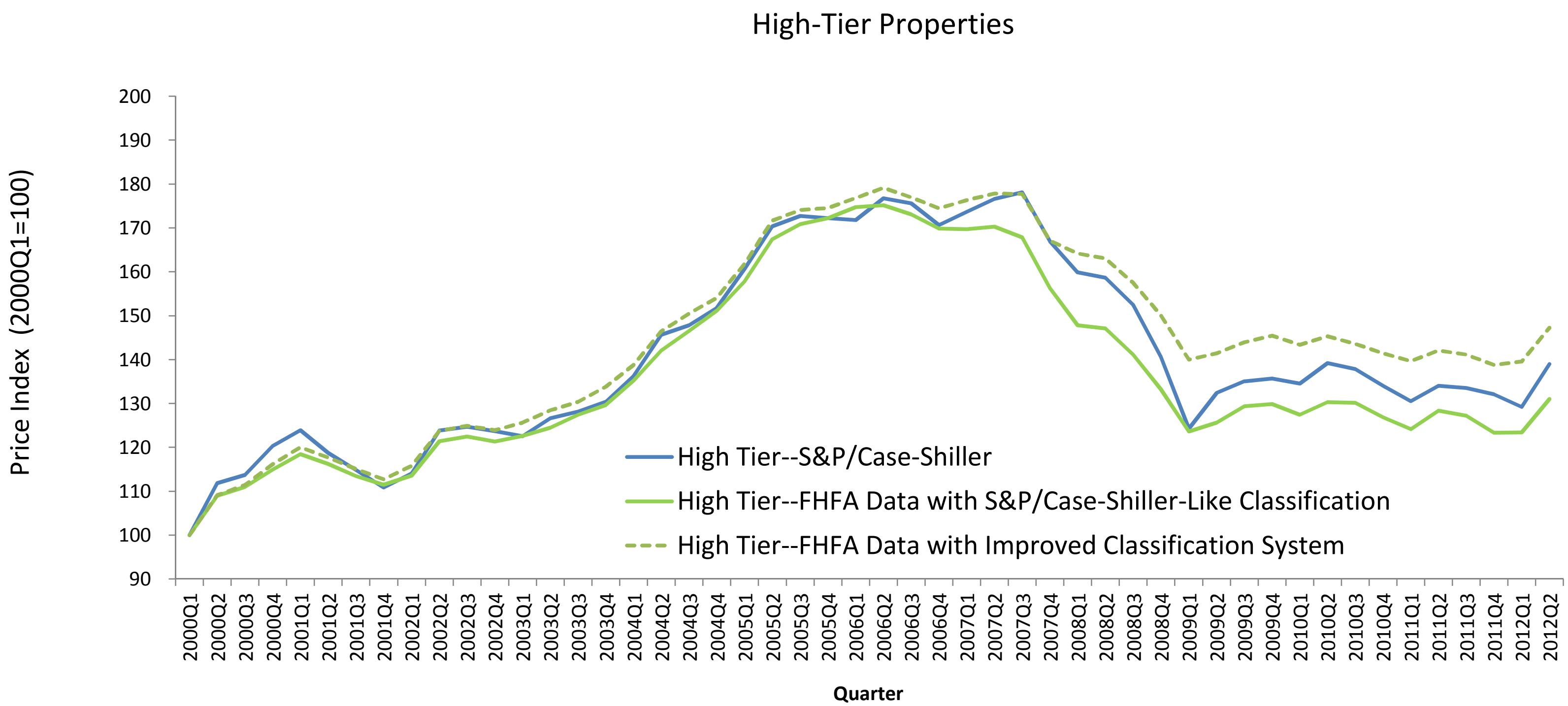


Table 1: Tier-Specific Price Changes Calculated under Alternative Tiering Approaches

(California; FHFA "Expanded-Data" House Price Index)

	Boom Appreciation (2000Q1-Peak)		**Bust Depreciation (Peak-2012Q2)**		**Aggregate Change (2000Q1-2012Q2)**	
	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home
Low Tier	197.9%	182.2%	-55.3%	-58.3%	33.2%	17.7%
Middle Tier	159.8%	162.9%	-49.5%	-47.7%	31.1%	37.6%
High Tier	111.8%	118.4%	-36.3%	-30.8%	34.9%	51.0%

Source: "Expanded-Data" HPI formed using mortgage-level data from the Enterprises and FHA, as well as county recorder data licensed from DataQuick Information Systems.

Table 2: Tier-Specific Price Changes Calculated under Alternative Tiering Approaches
(San Francisco; FHFA "Expanded-Data" House Price Index)

	Boom Appreciation (2000Q1-Peak)		Bust Depreciation (Peak-2012Q2)		Aggregate Change (2000Q1-2012Q2)	
	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home
Low Tier	156.3%	141.4%	-52.5%	-56.4%	21.6%	5.2%
Middle Tier	109.4%	111.6%	-41.5%	-38.3%	22.6%	30.6%
High Tier	75.2%	79.1%	-25.2%	-17.8%	31.0%	47.2%

Source: "Expanded-Data" HPI formed using mortgage-level data from the Enterprises and FHA, as well as county recorder data licensed from DataQuick Information Systems.

Table 3: Tier-Specific Price Changes Calculated under Alternative Tiering Approaches
(San Diego; FHFA "Expanded-Data" House Price Index)

	Boom Appreciation (2000Q1-Peak)		**Bust Depreciation (Peak-2012Q2)**		**Aggregate Change (2000Q1-2012Q2)**	
	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home	Tiering Based on First Transaction	Tiering Based on All Normalized Prices Available for Home
Low Tier	178.1%	171.5%	-43.1%	-46.5%	58.4%	45.2%
Middle Tier	144.1%	143.9%	-38.7%	-37.5%	49.7%	52.3%
High Tier	112.1%	114.8%	-32.1%	-27.8%	44.1%	55.1%

Source: "Expanded-Data" HPI formed using mortgage-level data from the Enterprises and FHA, as well as county recorder data licensed from DataQuick Information Systems.