# A Guide to Aggregate House Price Measures

*By Jordan Rappaport*

I n recent years, the United States, like many other industrialized nations, has experienced wide swings in the growth rate of housing prices. Understanding these price changes is important for a number of reasons. Housing serves as a major source of individual wealth. Hence, changes in its value may influence consumer spending and saving decisions, in turn, affecting overall economic activity. More narrowly, changes in housing prices both impact and reflect the health of the residential investment sector, a major source of employment. Further, house prices are the key determinant of housing affordability, an important public policy goal in many countries.

To understand the behavior of housing prices and their influence on the economy, it is crucial to have an accurate measure of aggregate housing prices. In practice, however, it is difficult to develop such a measure. Housing is an extremely heterogeneous good, and houses are sold only infrequently. Heterogeneity makes it difficult to distinguish between aggregate and individual price variations. The infrequency of sales implies that, in any time period, prices are not observed for most houses.

*Jordan Rappaport is a senior economist at the Federal Reserve Bank of Kansas City. Martina Chura, a research associate at the bank, helped prepare the article. This article is on the bank's website at **www.KansasCityFed.org**.*

In the face of such challenges, three methodologies have been developed to measure the aggregate price of housing. The first methodology simply takes an average over all observed prices, with no attempt to control for heterogeneity. The second looks at repeat sales of the same property. The third treats a house as a bundle of attributes, each with its own price that changes over time.

This article provides an overview of the three methodologies for pricing housing and a detailed guide to the major house price indices used by housing analysts. The analysis suggests there is no one "best" measure of housing prices. Each of the three methodologies has conceptual advantages and disadvantages, and the empirical house price indices have practical advantages and disadvantages as well. Which is best depends on the question being addressed.

The first section of the article examines why it is so difficult to measure the aggregate price of housing and how each of the three methodologies addresses these problems. The following three sections describe each methodology in more depth and describe some leading house price series based on each. The final section briefly compares the behavior of a representative series of each methodology over the period 1990 through 2006 and then provides examples of problems for which one or another representative series is likely to be most appropriate.

## I.    DIFFICULTIES MEASURING THE AGGREGATE PRICE OF HOUSING

The two main problems measuring the price of housing are heterogeneity and the infrequency of sales.[1] This section explains how the interaction of these two problems interferes with measurement. Next, it introduces each of the three methodologies that try to overcome these problems.

### Measurement problems

The first of the two measurement problems is the tremendous *heterogeneity* among houses. No two houses are the same. At the very least, they differ in location. They may differ in neighborhood, city, or metro area. Even the difference of a few hundred feet can have an appreciable price effect. Obviously, so too will other attributes, ranging from the

number of bedrooms and bathrooms to building style and state of repair. Naturally, observed differences in characteristics between two houses will be reflected in differences in price.

The specific combination of attributes, both locational and physical, associated with any particular house can be thought of as corresponding to that house's "quality." As is intuitive, quality captures the grade of workmanship and materials within a house. But it also is meant to capture virtually every other variation in attributes. So, for example, a two-bedroom house is of higher quality than a one-bedroom house (all else equal). Similarly, a house in a desirable location is of higher quality than a house in an undesirable location (all else equal).[2] Unfortunately, the quality associated with any specific house is not directly observable. Otherwise, aggregate price could be measured as the price of a house of predetermined quality.

One consequence of heterogeneity is that the average quality of the U.S. housing stock changes over time. In particular, the average quality of U.S. houses has been increasing. Newly constructed homes are, on average, larger than existing ones. The resulting increase in average quality implies an increase in average price. This would be true even if the prices of all existing houses had remained unchanged. Clearly, it would be undesirable to have the aggregate housing price reflect such a change in quality.

The second of the two measurement problems is the *infrequency of sales*. Because transactions on any specific house occur relatively infrequently, it is hard to know the amount at which a specific house will transact today. Sales amounts for similar houses offer some guidance. But, at the very least, differing locations affect a house's quality.

Combined, the infrequency of sales and heterogeneity make it difficult to find a representative sample of home prices with which to estimate aggregate prices. The main way to get sample prices for a given time period is to look at homes that are sold in that period. But the quality of the homes that are sold may systematically differ from the quality of the overall housing stock. For example, it may be that one period's sales are disproportionately skewed toward low-quality houses thereby biasing down any estimate of aggregate prices. A different period's sales may be skewed toward high-quality houses. Alternatively, it

may be that owners of houses that have declined in price hesitate to sell, whereas owners of houses that have appreciated are anxious to sell. In this case, estimates of aggregate price appreciation will be biased upward.[3]

*Three methodologies*

There are three distinct approaches to measuring aggregate house prices. Each one deals differently with the issues of heterogeneity and infrequent sales. One approach takes a simple average of all house prices observed in a period—usually a mean or median. Doing so essentially ignores the problems raised by heterogeneity and infrequent sales. The benefit is that price series employing this average methodology can often summarize an immense number of transactions on a timely basis.

A second approach—the repeat sales methodology—focuses on houses that have sold more than once. So long as the quality of the houses has remained unchanged, their rate of price appreciation is expected to be the same as the rate of aggregate house price appreciation. Price series employing the repeat sales methodology do a very good job of controlling for heterogeneity, while providing aggregate price estimates for numerous U.S. geographies.

A third approach—the hedonic methodology—uses statistical techniques to control for differences in quality. In particular, correlations between the sale price of homes and their attributes are used to estimate "prices" for various attributes, which are then used to calculate the sum total price of a representative bundle of attributes. Unfortunately, properly implementing the hedonic approach requires more detailed attribute data than are typically available. Nevertheless, a leading series employing the hedonic approach does an excellent job pricing an approximately constant-physical-quality new house over time.

## II.   AVERAGE PRICE MEASURES

The average methodology represents the simplest approach to measuring the aggregate price of housing. It simply measures the average of all *observed* housing prices. Typically, house prices are observed either due to a sales or a refinancing.

The average methodology essentially ignores the problems of heterogeneity and infrequent sales. Little or no attempt is made to assure the sample of houses whose price is being averaged is representative of the housing stock more generally. Nor is any attempt made that the sample remain comparable over time. Instead, it is hoped that with a very large number of transactions, the sample composition of houses will be sufficiently similar across time to give a reasonably accurate gauge of how the aggregate house price level has changed. Indeed, the average price methodology estimates the average price of the housing stock in its entirety. For example, the median of a large representative sample of house prices implicitly estimates the median of all house prices.

However, the average price methodology has some undesirable consequences that follow from the continual change in the housing stock's average quality. As described in the previous section, the construction of ever-larger homes would drive the average house price up even if all existing home prices remained unchanged. Ideally, an aggregate housing price should not reflect this aggregate composition effect.

A separate undesirable sample composition effect arises from the interaction of heterogeneity and the infrequency of sales. In any particular period, the sample of houses that sells may not be representative of the whole. The associated variations in average sample quality from period to period will suggest aggregate price changes that would not be implied by a truly representative sample.

Notwithstanding these problems, a huge advantage of the average methodology is its simplicity. As a result, sample sizes for estimates can be extremely large. And measures of U.S. housing prices as a whole, as well as for each of the four Census Bureau regions (Northeast, Midwest, South, and West) are available on a monthly basis.

Two leading measures of aggregate home prices employ the average methodology. The most widely cited, published by the National Association of Realtors (NAR), gives the median price of existing home sales. The Census Bureau also publishes a series that gives the median price of new home sales.[4]

*The National Association of Realtors existing home median value*

The NAR represents real estate brokers, property managers, appraisers, and other real estate professionals across the nation. NAR encompasses more than 1,700 local associations and boards. Many of these local boards also govern multiple listing services (MLSs) that help match house sellers and buyers. Each month, NAR surveys a fixed subset of its associations, boards, and MLSs that account for approximately 30 to 40 percent of all existing single-family home transactions. This sample group accounts for, on average, over 150,000 transactions per month. Existing single-family homes are attached or detached houses that are either currently or previously occupied. They exclude mobile homes, condominiums, and cooperative apartments.[5]

For each of the four Census Bureau regions, a median price is calculated based on the reported transactions from the sample. (The median price is the one at which half of the transactions are higher and half are lower.) A national median is then determined as a weighted average of the regional medians. The weights are the number of single-family units as counted in the 2000 decennial census. This regional weighting helps to limit the effect of shifts in the composition of the NAR sample. For example, an increase in the number of sales in the expensive Northeast relative to sales in the inexpensive Midwest will not affect the national median.

Nevertheless, the NAR median value is subject to considerable short-term volatility due to compositional changes. Within the four regions, any increase in the pace of sales of high-priced units relative to low-priced units will increase the regional median and hence the national median. This compositional effect is endemic to the average methodology.

The black line in Chart 1 shows the monthly time series of NAR median values. The estimated median transaction value, in current dollars, rises from $94,000 in December 1990 to $222,000 in December 2006. This 135 percent increase compares to a 51 percent increase in the Consumer Price Index over the same period.

Notice that even during the long secular rise in NAR median price from 1990 through 2004, there are numerous downward segments. In other words, even while the general trend in the estimated aggregate price was upward, the estimates also suggest there were short periods of aggregate price decline. A large portion of these declines is accounted for

*Chart 1*
## NAR MEDIAN PURCHASE PRICE, EXISTING HOMES

Current dollars

[Chart showing NAR median value (Existing homes) in current dollars, Y-axis from 50,000 to 300,000, X-axis from Dec-90 to Dec-06]
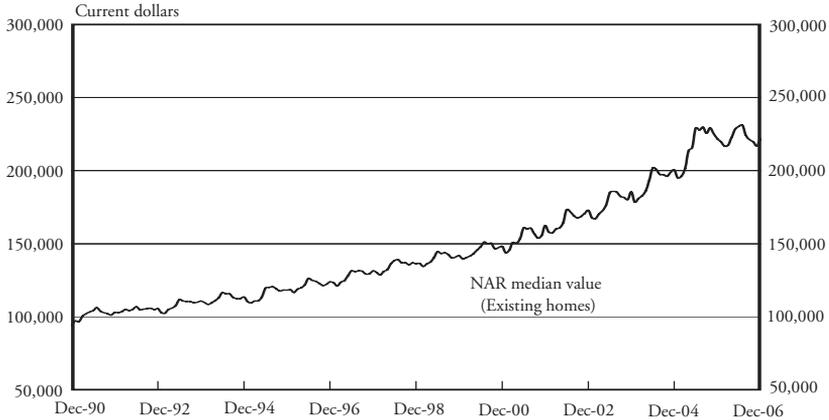
Source: National Association of Realtors

by seasonal fluctuations. But even after seasonally adjusting the data, downward segments remain. While it is possible that the "true" aggregate price declined (that is, the median price of all existing houses), a more likely explanation is that downward segments reflect some compositional shift in the sample group of sales.

One of the great benefits of the NAR series is the timeliness of its publication. While the aggregate measures based on the two other methodologies are only available on a quarterly basis, the NAR series is published monthly, both for the nation as a whole as well as for each of the four Census regions.[6]

*Census Bureau median value of new homes*

A second measure of average house prices focuses on new homes. The U.S. Census Bureau conducts a monthly survey of residential construction activity to estimate the rate and price of new home sales. For a representative sample of localities, the Census Bureau randomly samples single-family home permits. For each permit in the sample, the Bureau tracks when a deposit is taken or a contract is signed for the purchase of the home. Once either of these occurs, the Bureau considers the home to

*Chart 2*

## CENSUS BUREAU MEDIAN PURCHASE PRICE, NEW HOME PURCHASES
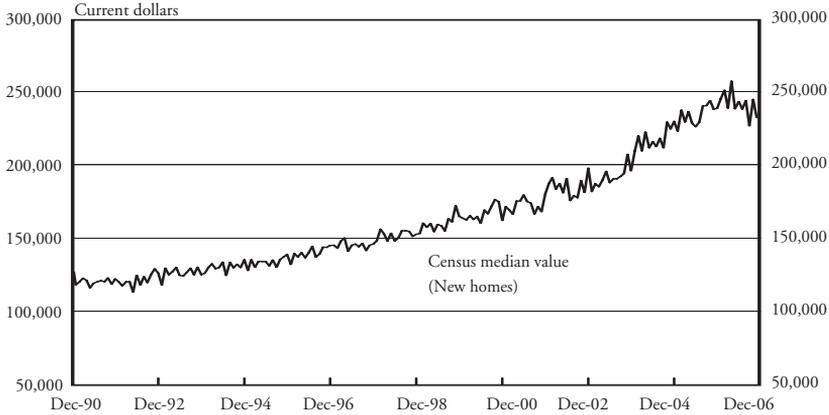


Source: U.S. Census Bureau

be sold and determines the transaction price.[7] From its sample of observed sales, the Bureau calculates a national median price.[8] This, in turn, implicitly estimates the national median price of all new housing units.

Chart 2 depicts the Bureau's estimated median value of new home sales from 1990 through 2006. It rises from $127,000 in December 1990 to $235,000 in December 2006, an increase of 85 percent. Like the NAR time series, the Census Bureau median values are upward sloping in the long term but contain numerous short-term downward-sloping portions. Again, it seems likely that these capture, in part, compositional effects arising from the sales sample rather than any true decline in aggregate price.

Also, like the NAR time series, the Bureau's median value of new homes is published on a monthly basis. Hence, it is a timely measure of aggregate prices compared to measures based on the repeat sale and hedonic methodologies, which are published only on a quarterly basis. The Bureau's median value of new home measure is not available for any U.S. geography other than the nation.

The median value measures of NAR and the Census Bureau are excellent estimates of the typical expenditure required to purchase housing, either existing or new. But because of compositional problems—both for

*Chart 3*
## GROWTH OF MEDIAN PRICES
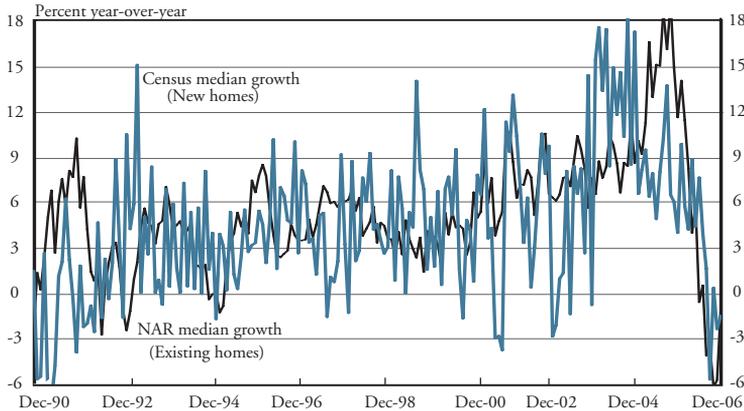


Source: National Association of Realtors, U.S. Census Bureau

the housing stock as a whole and, more prominently, of the sample for which prices are observed—they are less helpful for estimating the typical rate of house price appreciation. For instance, year-over-year growth rates of both series are characterized by high volatility, which is probably due to measurement issues rather than fundamental fluctuations (Chart 3).

To estimate typical price appreciation, it is preferable instead to look at repeat transactions of the same house.

## III.  REPEAT SALES PRICE MEASURES

The main problem with the average methodology is its inability to control for the changing quality of the houses in its price sample. The straightforward idea motivating the repeat sales methodology is that a house's quality remains approximately the same over time. If this is indeed so, then any observed change in a house's price must either be due to a change in aggregate prices or to some random "noise." Looking at price changes across a large number of houses filters out this noise and thereby estimates the path of aggregate prices.[9]

The most obvious problem with the repeat sales methodology is the constant-quality  requirement for houses that are included in the analysis. In fact, the quality of most houses  changes over time. On the one

*Chart 4*
## PER-UNIT EXPENDITURES FOR
## RESIDENTIAL IMPROVEMENTS
*All Owner-Occupied Properties*
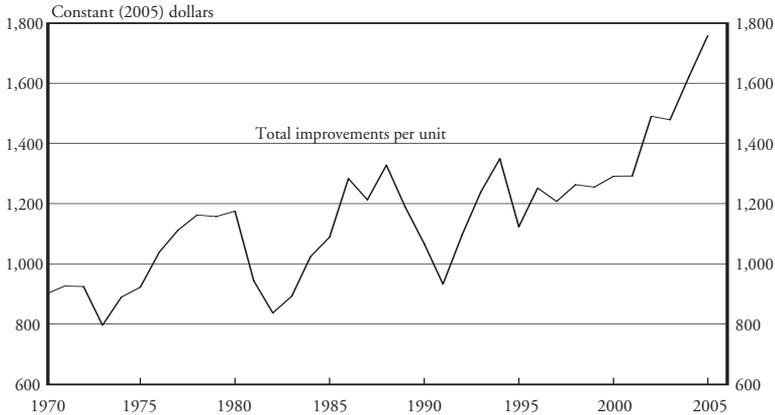


Source: U.S. Census Bureau

hand, houses age and can become rundown. In assuming that quality remains constant, the repeat sales methodology may thus underestimate the appreciation of the aggregate price of housing (Harding, Rosenthal, and Sirmans).

On the other hand, numerous owners devote considerable time and money on home improvement. To the extent that such efforts merely maintain the condition of their home, the repeat sales methodology will be accurate. But to the extent that they increase the quality of their home,  the repeat sales methodology may overestimate the rate of appreciation. Estimates for the 1970s and 1980s suggest that home improvement indeed biased upward growth rates based on repeat sales by ½ to 1 percent per year (Abraham and Schauman; Peek and Wilcox). As shown in Chart 4, expenditures on home improvement per unit were somewhat higher in the 1990s than in the 1980s and then spiked beginning in 2002. Hence, there is reason to suspect that the upward bias in repeat sales indices has become even larger.

A different problem with the repeat sales methodology, stemming from the infrequency of sales, is that it is subject to "transaction bias." Homes that are repeatedly sold may not be a representative sample of houses more generally (Case and Quigley; Hwang and Quigley). For one Hawaiian sample, houses sold at least twice were found to be con-

siderably more expensive than houses sold only once (Case and Quigley). Another study found that houses that transacted more frequently appreciated more rapidly (Case, Pollakowski, and Wachter). Still another study found that transaction bias caused price gains to be overstated during economic expansions and understated during declines (Gatzlaff and Haurin).

In addition to these two problems, the repeat sales methodology has several generic limitations and drawbacks. First, the repeat sales methodology can only estimate an index of the price level rather than the price level itself. The reason has to do with the statistical properties of the underlying regression equation.[10] Hence, for understanding, say, the affordability of housing, the repeat sales methodology is not helpful.

A second limitation is that the number of observed repeat transactions is small compared to the total number of sales transactions (which is used in the average price methodology). In one study of house price appreciation in four metro areas from 1970 through 1986, the number of usable repeat transactions was just 4 percent of total observed transactions  (Case and Shiller).[11] In other words, the usable number of repeat sales was just 4 percent the total number of sales. Other studies have found somewhat higher repeat sales shares: 11 percent (Case and Quigley) and 38 percent (Hwang and Quigley). Even the largest of these represents an important loss of precision in estimating a housing price level.

A third limitation is that a repeat sales index is subject to continual revision. That is, an initial estimate of the rate of house price appreciation between any two periods will continually be revised. The most intuitive explanation follows from houses whose "initial" sale is in the reporting period. The prices of such houses are not included in the estimation until a "repeat" sale occurs. Only then does the price appreciation of such houses affect the estimate of the initial reporting period's aggregate index level. As a quantitative example, consider the growth rate from 2005 Q2 to 2005 Q3 based on the Office of Federal Housing Enterprise Ovesight (OFHEO) House Price Index (HPI), one of the two repeat sales indices described below. Based on data through 2005 Q3, it was estimated to be 11.9 percent (annualized). But based on data through a year later, it was estimated to be 13.9 percent. This two-percentage-point change is on the high side for such revisions.

Moreover, revisions of growth rates over longer periods, such as a year or more, tend to be even smaller. Nevertheless, it needs to be recognized that a repeat sales index value is always tentative.

Notwithstanding these problems and limitations, the repeat sales methodology remains an excellent approach to estimating national house price appreciation. There are currently two main publicly available repeat sales indices of U.S. housing prices: the OFHEO HPI and the Standard and Poor's/Case-Shiller National Home Price Index.[12]

## OFHEO House Price Index

OFHEO is an independent agency within the Department of Housing and Urban Development. Its primary mission is to regulate the operations of the Federal National Mortgage Association (Fannie Mae) and the Federal Home Loan Mortgage Corporation (Freddie Mac). The latter two are government-sponsored enterprises that purchase home mortgages from lenders, which they then repackage into mortgage-backed securities to sell to investors or else hold themselves.

The OFHEO HPI uses price data collected by Fannie Mae and Freddie Mac each time they purchase a mortgage on a single-family home. Because Fannie and Freddie are such active participants in the U.S. mortgage market, the resulting database of property prices is extremely large. In 2005, Fannie and Freddie collectively purchased more than 4 million mortgages. The HPI time series, which begins in 1975, is based on more than 30 million paired mortgage transactions.

The HPI has several limitations in addition to those inherent to repeat-purchase indices in general. First, Fannie and Freddie's underwriting standards preclude them from making loans that exceed a "conforming" loan limit. This limit, which is reset annually, rose from $187,000 in 1990 to $253,000 in 2000 and $417,000 in 2007.[13] Price information on properties that are significantly more expensive than these limits may be contained in the Fannie and Freddie database so long as the mortgage itself conforms. But such expensive properties will more likely be financed with a nonconforming "jumbo" loan. Hence, high-priced homes will be underrepresented in the HPI sample.[14] A related consequence is that homes from especially expensive metro areas

are likely to be underrepresented. For example, NAR's estimates of the 2006 median home price in each of San Francisco, San Jose, and Anaheim exceeds $700,000.

Just as the HPI tends to underrepresent high-priced homes, it also tends to underrepresent low-priced homes. The reason is that Fannie and Freddie only purchase "conventional" loans, which exclude typically smaller-sized mortgages that are insured by the Veterans Administration and the Federal Housing Administration. Also excluded from the HPI are condominiums, co-ops, and other multifamily homes. These and the previous exclusions imply that  the HPI is best interpreted as  measuring aggregate price appreciation for a broad middle segment of the U.S. stock of single-family homes.

In contrast to these limitations of exclusion, a potential limitation of inclusion is that the HPI is based on mortgages issued both for sales and for refinancings. For the latter, the recorded price of a house is its appraised value, which almost certainly differs from the price at which the house would sell were it on the market. More specifically, some studies suggest that appraisals systematically overestimate the value of homes, perhaps to win more work from mortgage brokers and realtors (Ferguson; Gwin, Ong, and Spieler). The benefit of including refinancing transactions is that it greatly increases the number of paired transactions. Paired transactions that include at least one refinancing account for approximately 85 percent of all paired transactions in the HPI database (Stephens and others). Hence, excluding them may increase the noisiness of estimated price appreciation.

Estimated price appreciation over long periods proves relatively insensitive to whether refinancings are included.[15] In addition to the HPI, the OFHEO publishes an alternative repeat-sales index based just on purchases of houses, excluding refinancings. In levels, the purchase-only index and the HPI are barely distinguishable from each other after December 1990. In growth rates, the HPI fluctuates somewhat more than the purchase-only index (Chart 5). This probably makes the purchase-only index a better measure of short-term price growth. One particular noticeable difference between the two time series is the HPI's much higher growth rate from the second half of 2004 through mid-2006. This faster growth

*Chart 5*
## OFHEO HPI AND PURCHASE-ONLY INDEX GROWTH



Percent year-over-year

House price index

Purchase-only index

Source: Office of Federal Housing Enterprise Oversight

is consistent with overly optimistic appraisals accompanying refinancings.[16] Notwithstanding such differences, the contemporaneous correlation between the two growth time series is 0.94.

A final note on the HPI: the national index is actually a weighted average of separate indices calculated for each of the nine Census Bureau divisions.[17] In other words, OFHEO pools all repeat transactions within each division to estimate a price index. These are then weighted by the division's number of single-family detached housing units from the 2000 decennial census to calculate a national index. This weighting scheme helps prevent the national index from becoming unduly influenced by markets with especially large numbers of transactions, but not necessarily a large number of actual houses.

OFHEO also publishes indices for each of the U.S. states and for most metropolitan areas. Like the HPI, they are available on a quarterly basis.

### S&P/Case-Shiller U.S. National Home Price Index

A second repeat sales index, published by Standard & Poor's, is known as the S&P/Case-Shiller U.S. National Home Price Index (SPCSI).[18] Its underlying data come from deed records of residential sales transactions. The SPCSI thus has greater scope than the HPI, which relies on mortgages purchased by Fannie Mae or Freddie Mac. As

with the various other measures, the SPCSI tracks the value just of single-family homes, which excludes condos and co-ops. Substantial effort is also devoted to exclude transactions that are not made at "arms length." For example, transactions in which the seller and purchaser have the same surname are excluded. The worry is that a non-arms-length transaction price will differ from a true market price. Similarly, substantial effort is made to exclude transactions following large changes in property attributes. For example, purchase-sale pairs that occur within six months of each other are excluded due to the possibility that a developer has quickly upgraded and sold a property.

Like the HPI, the SPCSI is constructed as a weighted average of repeat sales indices for each of the nine Census Bureau divisions. In contrast to the HPI, which weights each of the divisions by its number of housing units in 2000, the SPCSI weights them by the estimated aggregate value of housing in 2000. Doing so implies that Census divisions with more expensive housing have more sway in determining the national index. Moreover, each of the nine division indices is constructed using a variation of the repeat sales methodology that weights each repeat sales pair by the value of the initial transaction. This value weighting of observations similarly implies that price appreciation by an initially more expensive house exerts more influence on the division price index than does price appreciation by an initially less expensive house.

Value weighting is an important way in which the SPCSI differs from the HPI. It makes the SPCSI a good estimate of the investment return to holding a "basket" of housing that is representative of the nation. Since houses in some parts of the country cost more than in others, value appreciation by them contributes more to such an investment return compared to an equal-percentage value appreciation by houses in low-cost parts of the country.[19]

As a consequence of its value weighting, the SPCSI growth rate systematically differs from the HPI growth rate. In Chart 6, the solid teal line represents the year-over-year growth of the SPCSI. The black line represents year-over-year growth of the HPI. During the early 1990s, growth of the HPI exceeded that of the SPCSI. Indeed, SPCSI growth was negative from Q4 1990 through Q3 1991. This was a period when house prices were typically declining in many expensive metro areas such as New York, Boston, Los Angeles, and San Francisco. Because of its value weighting, the SPCSI will tend to more closely track house

*Chart 6*

## OFHEO HPI AND S&P/CASE-SHILLER NATIONAL INDEX GROWTH

Percent year-over-year

SPCSI national

OFHEO HPI
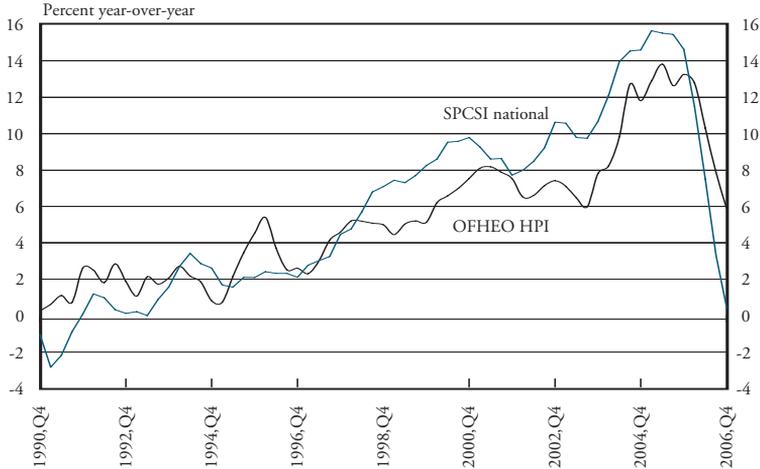
Source: Office of Federal Housing Enterprise Oversight, Standard & Poor's

prices in such metro areas than will an index that weights houses equally, as does the HPI. This would account for the negative growth of the SPCSI but positive growth of the HPI.

Conversely, SPCSI growth consistently exceeds HPI growth starting in 1998. Over the subsequent six years, it does so by as much as four percentage points. During this period, house prices were rising especially fast in more expensive metro areas, particularly in California.[20]

Overall, the OFHEO's HPI is a good estimate of the typical price appreciation of single-family houses, whereas the S&P/Case-Shiller index is a good estimate of the capital appreciation that would result from owning a representative sample of U.S. homes. For most uses, the former aggregate price measure is probably more relevant.[21] On the other hand, S&P/Case-Shiller dominates the HPI by accounting for all arms-length transaction, not just those securitized by Fannie Mae or Freddie Mac. Two methodological problems shared by both indices are that measured price appreciation may actually be capturing improvements and that the sample of houses with repeat sales may not be representative of houses more generally. In part to address such problems, researchers turn to hedonic techniques.

## IV.  HEDONIC TECHNIQUES

Instead of assuming that a house's quality remains constant over time, the hedonic statistical methodology explicitly estimates prices for the attributes that determine house quality. It can then "construct" and price a hypothetical constant-quality house, that is, one with the same attributes over time. By choosing the appropriate mix of attributes, this constant-quality house is taken to be representative of the aggregate housing stock.

One interpretation of the services that flow from any particular house is that they represent the sum total of the services that flow from each of its many attributes. In other words, a house's service may represent the sum of its bedroom services, bathroom services, kitchen services, lot services, location services, etc. If so, a house's price would approximately equal the sum total of the price times the quantity of each of its attributes. This interpretation implies a straightforward statistical regression that estimates attribute prices based on the correlations between observed house prices and house attributes.

To estimate an aggregate price of housing, the final step is to apply the estimated attribute prices to a set of attributes representative of the aggregate housing stock. Typically, the representative bundle comprises the estimated average quantity of each attribute in the housing stock in some base year. The price of the representative bundle becomes the estimate of aggregate prices. Note that because attribute prices may change over time at different rates, the composition of the representative bundle meaningfully affects the rate of aggregate price appreciation. Hence, so too does the choice of base year.[22]

The main drawback to this hedonic technique is that it requires a tremendous amount of data on house attributes. As already discussed, the number of attributes that affect a house's quality is extremely large. The subjective nature of some of these attributes makes them difficult to determine even for a single house. Obtaining the information on the attributes for the thousands of observations necessary to estimate corresponding prices is thus a substantial challenge.

*Census Constant Quality Index of New One-Family Homes Sold*

Because it is so difficult to obtain the necessary information, the only regularly published hedonic aggregate house price series is the Census Constant Quality Index of New One-Family Homes Sold (CCQI). As part of its monthly survey of residential construction activity described in Section II, the Census Bureau records the quantity of approximately 20 different attributes. On a quarterly basis, it runs a hedonic regression to estimate the individual price of 12 of these attributes.[23] It then calculates the price of a fictional house, using as the quantity of each attribute its arithmetic mean from a base year. As of 2007, the base year is 1996.

More specifically, the Census Bureau actually runs five separate regressions. For detached single-family homes, it runs one for each of the four census regions. And it runs a fifth for attached single-family homes (for example, town houses). This separation allows the estimated price of a given attribute to differ by region and house type. Thus, for example, the estimated price of a fireplace may be higher in the Northeast than in the South. Similarly, an extra bedroom may be more valuable in a detached home than in an attached one. For each of the five regressions, the Bureau calculates an average price level based on the 1996 base characteristics and the currently estimated attribute prices.

The five separately calculated price levels are then combined. Together, they form a single U.S. aggregate price level using weights based on the number of new single-family detached and attached units *sold* in 1996. The resulting weighting heavily emphasizes the South and West over the Northeast and Midwest, as the former were experiencing much more rapid new home sales. One consequence is that the CCQI is an especially unrepresentative measure of existing single-family home prices.[24]

Of course, the CCQI is not meant to measure existing single-family home prices but rather new home prices. Nevertheless, it is tempting to use it to do so, given that there is no hedonic measure of existing home prices. A second reason why it is a mistake to do so concerns the location within metropolitan areas of new development. New construction typically occurs at the outer edge of a dense metropolitan settlement, wherever that outer edge happens to be. To the extent that land is less

*Chat 7*
## CENSUS CONSTANT QUALITY INDEX HOUSE PRICE LEVEL
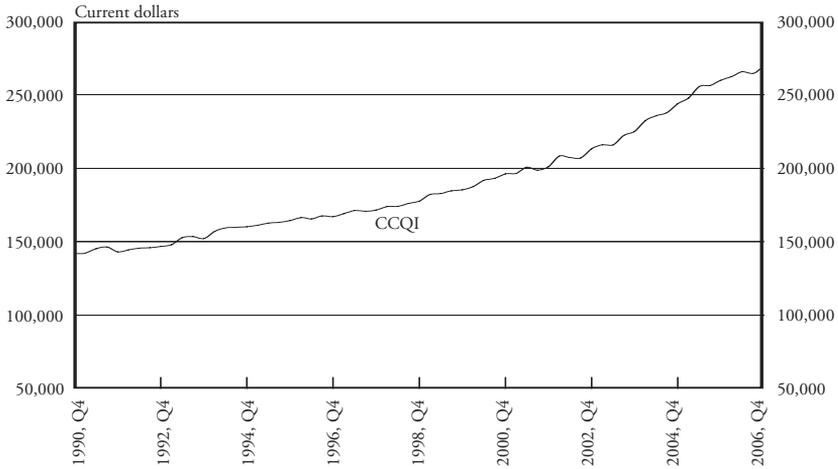


Source: U.S. Census Bureau

expensive at this outer edge, the constant-quality index will poorly measure the change in prices at a fixed location. That is, as a location's price rises, developers typically move outward to less-expensive locations.

Chart 7 shows the level of the CCQI from 1990 through 2006. As mentioned above, the representative bundle of characteristics is benchmarked to observed averages in 1996. The resulting 1996-quality house rose in price from $142,000 in Q4 1990 to $167,000 in Q4 1996 to $269,000 in Q4 2006. The estimated 89 percent aggregate price appreciation over this period is approximately the same as is estimated by the simple median price of new housing discussed in the second section. This is surprising. The median-priced new home reflects compositional changes over time. Since houses have been becoming larger, the median-priced new home should appreciate more quickly than a constant-quality new home. One possible explanation is the acceleration of construction in relatively low-cost cities such as Austin, Phoenix, and Raleigh-Durham. This would put stronger downward pressure on the median home price than on the constant-quality price, since the latter holds regional weights constant. During the 1980s, in contrast, the median price did indeed grow significantly quicker than did the constant-quality price.[25]

Overall, the CCQI of New One-Family Homes Sold does a good job of estimating the aggregate price for new homes at metro area fringes. But because of its regional weighting and the lack of control for location within a metro area, the CCQI is probably a poor measure of existing home prices. The next section, nevertheless, compares its time path with those of the NAR median home price and the OFHEO repeat sales index.

## V.    COMPARISON AMONG TYPES

This section directly compares examples of each of the three types of measure: a simple average of house prices, a repeat sales derived price, and a hedonic price. Beforehand, it briefly discusses some research comparing the methodologies themselves. Afterward, it concludes with some specific recommendations on choosing among the various house price series.

### *Research comparing methodologies*

A number of researchers have compared examples of each methodology. The general aim is to ascertain which of the methodologies best measures the typical appreciation of a constant-quality existing house. The only clear conclusion from this research is that when the number of repeat sales transactions is low, the repeat sales methodology does poorly.

The poor performance of the repeat sales methodology is documented by Meese and Wallace, who study price appreciation in Oakland and Fremont, California, during the 1970s and 1980s. They have price data for a sample group of houses that numbers well over 20,000. But only about 3,000 of these are repeat sales transactions.[26] The authors argue that the estimated price appreciation, using the repeat sales methodology, is implausibly steep. They attribute this to the biased nature of the repeat sales sample. They also argue that a few observations may unduly affect the estimate. In other words, idiosyncratic appreciation by a few houses may incorrectly be inferred as aggregate appreciation. Both the median and hedonic methodologies suggested a similar rate of house price appreciation. The authors conclude that the repeat sales methodology performs poorly when the number of observations is small.

Using municipal data, a different group of researchers argue that the repeat sales estimator performs poorly over short time periods but well over long ones (Clapp, Giaccotto, Tirtiroglu). The reason for the poor performance, as mentioned earlier, is the small number of repeat transactions. Moreover, such observations appeared not to be representative of houses more generally. On the other hand, over periods of three years and longer, the repeat sales sample appreciated at a rate similar to that of all houses. Thus, the repeat sales methodology may be a reasonably accurate way to measure long-term price appreciation.

The repeat sales methodology also holds up well when it has the same number of observations to work with as the other two. Crone and Voith study the price appreciation of approximately 14,000 homes in a rural Pennsylvania county during the 1970s and 1980s. All of the homes were sold at least twice. Half of the homes were used to estimate prices by each of the three methodologies, and the remaining half were used to test the accuracy of predicted prices. The prediction errors were by far the largest using the average methodology. More specifically, using median (or mean) prices to estimate the price appreciation of homes was far less accurate than using the repeat sales or the hedonic techniques. Between the latter two methodologies, there was no clear favorite. The repeat sales methodology produced smaller *average* prediction errors. But many of its individual prediction errors were especially large.

Overall, no clear consensus exists on which methodology is best. The failure to control for compositional shifts makes the average methodology unattractive in theory. But as is shown in the next section, in practice, national median prices have actually behaved similarly to repeat sales prices. The repeat sales methodology may perform poorly when the number of repeat transactions is low. But for estimating a national aggregate price, the number of repeat transactions is extremely high. From a theoretical point of view, the hedonic methodology is especially attractive because it explicitly controls for variations in housing quality. But the data to make a hedonic estimate of the national home price level is available only for new homes.

*Chart 8*
## NAR MEDIAN, OFHEO HPI, AND CCQI LEVELS
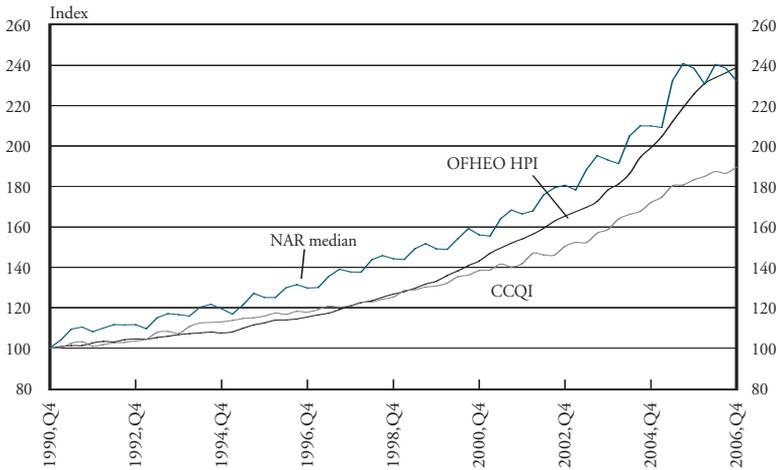


Source: NAR, OFHEO, U.S. Census Bureau

### *Comparing specific house price series*

One way to get a better sense of the relative strengths and weaknesses of the three methodologies is to compare aggregate price series based on each. Chart 8 shows indices of aggregate housing prices based on the NAR median price of existing homes (average methodology), the OFHEO HPI (repeat sales methodology), and the CCQI (hedonic methodology). The main disadvantage of this comparison set of indices is that the first two series are based on the price of existing homes, whereas the third is based on the price of new homes. Unfortunately, there is no hedonic time series of aggregate home prices based on existing homes.

Several important characteristics are evident in the chart. First is the relatively faster long-term growth of the NAR series throughout most of the 1990s. This is exactly what one would expect from a series based on the average methodology. Most likely, the faster growth captures the increasing quality of the U.S. housing stock.

Second is the relatively slower growth of the CCQI beginning in 1999. Such slower growth is exactly what would be expected theoretically. The CCQI does not control for the changing location of new homes and hence is likely to be biased downward.

Third is the faster growth of the HPI from 1999 forward. As just stated, theory suggests that the NAR aggregate price measure should grow the fastest. One possible explanation why this was not so is the HPI's upward bias from not controlling for home renovation, which was booming in the 1990s and especially after 2002. In addition, houses that have appreciated the most generally have the highest likelihood of being sold (and thus being included in the OFHEO database).[27] A second possible explanation is the increase in home ownership, which grew from 64 percent of households in 1990 to 69 percent in 2006. A large portion of this increase was by lower-income households that presumably demand lower-quality homes. Hence, there would be some corresponding downward shift in composition of the NAR sample.[28]

A fourth characteristic of the three series is the relatively higher short-term volatility of the NAR index.[29] Notwithstanding its long-term upward trend, the NAR index repeatedly spends short periods—typically one or two quarters—declining. This oscillating pattern primarily is accounted for by seasonality. Such seasonality, in turn, derives, in part, from regular shifts over the course of a year in the composition of houses sold. The CCQI is also characterized by moderate volatility, with its upward trend punctuated by a few periods of short-term decline. Only the HPI rises smoothly over the entire period shown.

The relative volatility of the three series is also evident in Chart 9, which shows year-over-year growth rates. The NAR and CCQI growth rates are characterized by numerous short-term up and down swings. Such volatility is not necessarily an undesirable property. After all, that may indeed be how the aggregate price of housing is behaving. But at least for the NAR series, it seems more likely that the fluctuations may be capturing short-term compositional shifts. Compared to these two series, the HPI, has lower short-run volatility. Of course it is possible that the HPI is missing actual aggregate volatility. But given the smooth behavior of nonhousing aggregate price series, such as CPI less food and energy, the relative smallness of the HPI's short-term fluctuations is probably a strength. One benefit is that the HPI will be the most effective among the indices in measuring short-term aggregate price changes, such as from quarter to quarter.

*Chart 9*
## NAR MEDIAN, OFHEO HPI, AND CCQI GROWTH RATES
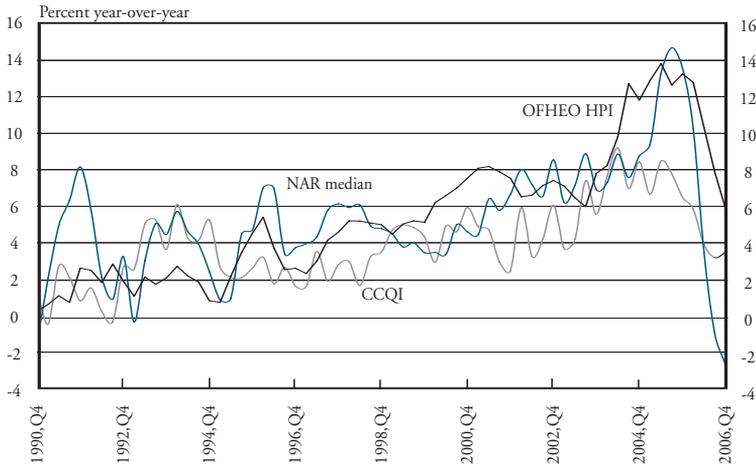


Source: NAR, OFHEO, U.S. Census Bureau

## *Choosing among the house price series*

More generally, which of the measures best estimates the aggregate price of housing? The answer very much depends on one's purpose.

One possible purpose seeks to estimate the *typical* increase in home-owners' wealth from an increase in house prices. To estimate such an aggregate price rise, the OFHEO HPI is probably best. As a repeat sales index, it does a reasonably good job of controlling for variations in houses' quality. And its low volatility suggests that even price changes measured over short periods are likely to reflect typical price changes for houses.[30] The main caveat is that the typical rate of price change is likely to be somewhat overstated, due both to the failure to control for home improvement and the unrepresentative nature of homes with repeat sales.

A second, related, purpose seeks to estimate the *aggregate* change in household net worth due to the increase in house prices. In other words, rather than estimating the typical gain, this purpose seeks to estimate the total gain. This is exactly the aggregate price estimated by the S&P/Case-Shiller index.

A third possible purpose is to gauge the health of the residential construction sector. Rising prices are an obvious spur to residential investment. On the one hand, developers have a greater incentive to build. On the other hand, owners of existing homes can extract equity and engage in home improvement. Here, both the CCQI and OFHEO HPI measures are appropriate. Most obviously, the CCQI gives the rate of price inflation of new homes against which developers will be competing. Similarly, the HPI gives the typical increase in existing home prices against which developers will also be competing. In addition, the OFHEO HPI will reflect the typical increase in home equity against which owners can borrow to fund home improvements.

A fourth possible purpose for estimating the aggregate price of housing is to gauge average affordability. For this, the NAR series is probably best. The short-term fluctuations of the NAR median price are largely irrelevant as they primarily affect the rate of price growth rather than its level. The OFHEO HPI is definitely not appropriate as it only estimates an index level, not a price level. And the CCQI is probably inappropriate unless one is interested in the affordability of a 1996-quality new house. On the other hand, the Census Bureau's median new home price would also be relevant. It addresses the affordability of the typical new home.

Of course, there are a multitude of other purposes for which aggregate house price measures are required. Understanding the strengths and weaknesses of the methodologies underlying the available series, as well as the practical details of the specific series themselves should help choosing one appropriate to the purpose at hand. With this goal in mind, Table 1 provides a summary of the house price series discussed in this article.

## IV. SUMMARY AND CONCLUSIONS

Numerous measures exist of the aggregate price of U.S. housing. Often they suggest very different rates of price appreciation. Which rate is correct can have important implications for a number of policy issues, such as consumer spending and saving decisions, the strength of the residential investment sector, and housing affordability.

*Table 1*
## SUMMARY OF HOUSE PRICE MEASURES

| Measure | Methodology | Home Type | Frequency | Additional Geographies | Comment | Website |
|---|---|---|---|---|---|---|
| NAR median home price | Average | Existing | Monthly | Regions; metro areas (quarterly) | Wide coverage | www.realtor.org/Research.nsf/Pages/EHSdata |
| Census median home price | Average | New | Monthly | | High volatility | www.census.gov/const/www/newressalesindex_excel.html |
| OFHEO house price index | Repeat sales | Existing | Quarterly | Divisions, states, metro areas | Probably biased upward | www.ofheo.gov/HPI.asp |
| OFHEO purchase-only index | Repeat sales | Existing | Quarterly | | Excludes refinancings; less volatile than HPI | www.ofheo.gov/HPI.asp |
| S&P/Case-Shiller National Home Price Index | Repeat sales | Existing | Quarterly | 20 metro areas | Value weighted | www.homeprice.standardandpoors.com |
| Census Bureau Constant Quality Index | Hedonic | New | Quarterly | Regions | Regional weighting emphasizes South and West | www.census.gov/const/www/constpriceindex.html |

The numerous price measures reflect the difficulty in pricing housing. The key reasons for this difficulty include the heterogeneity of houses and their infrequent sales. Houses vary in quality, both in a cross section and over time.

The various price measures fall into one of three methodologies. The first methodology simply takes an average of all observed prices, with no attempt to control for heterogeneity or changing composition. The second methodology looks at repeat sales of the same property. The third methodology treats a house as a bundle of attributes, each with its own price that changes over time.

Each of the methodologies has strengths and weaknesses. The average methodology is simple and has the lowest data requirements. But it misses both short-term fluctuations in the composition of housing as well as long-term changes in the quality of the typical U.S. house. The repeat sales methodology does an excellent job of controlling for cross-sectional variations in attributes among houses, but it has difficulty controlling for changes in house quality that occur between sales. Moreover, houses that sell repeatedly may not be representative of houses more generally. The hedonic methodology in theory does an excellent job of controlling for variations in house quality. But the tremendous data required to estimate a hedonic price index limit doing so to new homes.

There is no "best" methodology or price series. Rather, each may be best matched to one or another question. Given the multitude of questions that center on housing, it is therefore helpful to have so many house price measures.

# ENDNOTES

[1]A third, frequently cited problem is that recorded sales prices do not reflect any "give backs" by the owner to the buyer. Especially in weak markets, actual prices will tend to be slightly below nominal ones.

[2]A house of high quality can equivalently be thought of as a house that yields a high *quantity* of housing services.

[3]The sample composition effects stem from the interaction of heterogeneity and infrequency of sales. With heterogeneity but very frequent sales, a representative large group of homes could be sampled each period to estimate an aggregate price. With homogeneity but infrequent sales, whichever homes happened to sell in a period would be sufficiently representative of homes in general to estimate an aggregate price.

[4]A third series, published by the Federal Housing Finance Board (*www.fhfb.gov/ Default.aspx?Page=53*) gives the mean value for each new and existing house on a monthly basis. The sources are mortgages written by member banks. The NAR series is generally preferred to the FHFB one because of its larger sample size. Moreover, the FHFB series excludes houses for which the mortgage is government insured as well as those with mortgages that are not fully amortizing. Nevertheless, the FHFB series is important because it establishes the upper limit for "conforming" loans, which are discussed in the text below.

[5]NAR also publishes a series on the median price of condos and co-ops as well as a series on the combined median price of condos, co-ops, and single-family homes.

[6]NAR additionally publishes median prices for metropolitan areas on a quarterly basis.

[7]In contrast, NAR considers a home as sold when the transaction closes; a disadvantage of the Census Bureau approach is that a significant portion of "sales" never close.

[8]Unlike the NAR series as well as the Census Bureau constant quality series described in Section IV, there is no attempt to hold constant the regional mix of sales.

[9]More specifically, a statistical regression can find a best-fit price path that minimizes the sum of differences between the rates of individual houses' price appreciation and the rate of aggregate price appreciation. Each paired sale serves as a single observation for the regression. The regression's dependent variable is the change in log price from initial sale to repeat sale. The regression's right-hand side is made up of dummy variables for each period except the first. For any particular observation, the dummy variables each take the value of zero except for the dummies corresponding to the initial sale and the repeat sale. These take the values -1 and 1, respectively. The estimated coefficients on the dummies give an index of the aggregate price path (Case and Shiller, 1987).

[10]One of the time-specific dummy variables needs to be dropped to avoid colinearity. Usually, the dummy for the first period is dropped. Doing so implies a logarithmic price level of zero in the first period, or equivalently, an exponentiated price level of 1.

[11]In addition to single sale properties, properties that were thought to have undergone substantial changes in condition were excluded from the sample.

[12]A third index, the Freddie Mac Conventional Mortgage Home Price Index, is nearly identical to the OFHEO HPI and so is not discussed.

[13]Changes in the conforming loan limit are determined by aggregate house price growth as estimated by the FHFB average price series discussed in endnote 4. The conforming loan limit is 50 percent higher in Hawaii, Alaska, the U.S. Virgin Islands, and Guam.

[14]Fannie's and Freddie's relatively strict underwriting standards also suggest that few, if any, of the mortgages they purchase will not be fully amortizing.

[15]For metro area indices, results are likely to be much more sensitive to the exclusion of refinancings. The reason is that the repeat sales methodology performs poorly when the number of transactions is low (Meese and Wallace).

[16]Abraham and Schauman compare repeat sales based on only refinancings versus on only arms-length transactions. A price index based on the former appreciates at approximately ½ a percentage point per year faster than does an index based on the latter.

[17]The nine census divisions are New England, Middle Atlantic, South Atlantic, East North Central, West North Central, East South Central, West South Central, Mountain, and Pacific.

[18]Standard & Poors also publishes, on a monthly basis, repeat sales indices for each of 20 major metropolitan areas. Alongside these, it publishes one index that summarizes the price appreciation in ten of the metro areas and another index that summarizes the price appreciation in all 20 metro areas.

[19]The SPCSI's value weighting makes it analogous to a capitalization-weighted stock index such as the S&P 500. Equivalently, the SPCSI is estimating the approximate appreciation of an arithmetic average of home prices. But in contrast to the average methodology, it controls for changes in quality.

[20]The SPCSI assigns a 22 percent weight to the Pacific Division, which includes California; the HPI assigns it a 14 percent weight.

[21]Measuring investment returns is the explicit purpose of the SPCSI. Consistent with this, futures contracts and options are traded on analogous value-weighted repeat sales indices published by Standard & Poor's for each of ten large U.S. metropolitan areas.

[22]The hedonic methodology is typically used to construct a Laspeyres index. Such an index takes representative attribute quantities from an initial period to measure prices in subsequent periods. A Laspeyres index tends to overstate price appreciation. Attribute prices may change at different rates from each other. As one attribute becomes more expensive, it is natural for buyers to look for houses that have a lower quantity of it but a higher quantity of some other, less expensive attribute. By doing so, they may be able to find a house yielding the same quantity of housing services as the representative bundle house but at a lower cost.

[23]The dozen attributes used in the hedonic regression are finished area square footage, geographic location (that is, state or Census division), inside/outside a metro area, number of bedrooms, number of bathrooms, number of fireplaces, type of parking facility, type of basement (finished/unfinished/none), presence of a deck, construction method, primary exterior wall material, and the types of heating and air conditioning systems. Some additional attributes recorded in the Census Bureau's *Survey of Construction* include lot size, number of floors, presence of a porch, and presence of a patio.

[24]For example, the West's 27 percent weighting in the CCQI would be just 19 percent based on existing single family homes as counted in the 2000 decennial census. And the Northeast's 6 percent share in the CCQI would rise to 15 percent based on existing single-family homes.

[25]The price level of the 1996-quality house is well above that of the median new house in 1996 and thereafter. This reflects that the representative bundle of new house attributes yielded more housing services than were associated with the median-priced house. More simply, the 1996-quality house was of higher quality than the median-priced new house. On the other hand, the mean-priced new home in 1996 was priced approximately the same as the constant-quality home.

[26]The repeat sales transactions included in the study were only those for which it was believed that no attribute changes had occurred.

[27]An additional possibility is that middle-priced properties, which constitute the bulk of the OFHEO database, may have grown quicker than either very low or very high priced properties, which are mostly excluded from the OFHEO database. However, the faster growth of the SPCSI series compared to OFHEO series suggests that higher-priced houses appreciated quicker than did lower-priced ones.

[28]Data on home ownership is available from *www.census.gov/hhes/ www/housing/ hvs/hvs.html*.

[29]In order to match the frequency of the other two series, the NAR index is calculated here as the quarterly average of monthly values.

[30]This desirability of low volatility suggests that the OFHEO purchase-only index might be even better than the OFHEO HPI.

## REFERENCES

Abraham, Jesse, and William S. Schauman. 1991. "New Evidence on Home Prices from Freddie Mac Repeat Sales," *AREUEA Journal*, vol. 19, no. 3, pp. 333-52.

Case, Bradford, and John M. Quigley. 1991. "The Dynamics of Real Estate Prices," *The Review of Economics and Statistics*, vol. 73, no. 1, February, pp. 50-58.

Case, Bradford, Henry O. Pollakowski, and Susan M. Wachter. 1997. "Frequency of Transaction and House Price Modeling," *Journal of Real Estate Finance and Economics*, vol. 14, pp. 173-87.

_____. 1991. "On Choosing Among House Price Index Methodologies," *AREUEA Journal*, vol. 19, no. 3, pp. 286-307.

Case, Karl E., and Robert J. Shiller. 1987. "Prices of Single-Family Homes Since 1970: New Indexes for Four Cities," Federal Reserve Bank of Boston, *New England Economic Review*, September/October, pp. 45-56.

Federal Financial Institutions Examination Council. 2006. "HMDA National Aggregate Report 2005," National Summary Table A1.

Ferguson, Jerry T. 1988. "After-Sale Evaluations: Appraisals Or Justifications," *Journal of Real Estate Research*, vol. 3, no. 1, pp. 19-26.

Gatzlaff, Dean H., and Donald R. Haurin. 1997. "Sample Selection Bias and Repeat-Sales Index Estimates," *Journal of Real Estate Finance and Economics*, vol. 14, pp. 33-50.

Gwin, Carl R., Seow E. Ong, and Andrew C. Spieler. 2006. "Real Estate Appraisal and Transaction Price: An Empirical Evaluation of Alternative Theories," *Journal of Housing Research*, vol. 15, iss. 1, pp. 29-39.

Harding, John P., 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*, forthcoming.

Hwang, Min, and John M. Quigley. 2004. "Selectivity, Quality Adjustment and Mean Reversion in the Measurement of House Values," *Journal of Real Estate Finance and Economics*, vol. 28, nos. 2/3, pp. 161-78.

Krainer, John. 2006. "Mortgage Innovation and Consumer Choice," Federal Reserve Bank of San Francisco, *FRBSF Economic Letter*, no. 38, December 29.

Meese, Richard A., and Nancy E. Wallace. 1997. "The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic-Regression, and Hybrid Approaches," *Journal of Real Estate Finance and Economics*, vol. 14, pp. 51-73.

Peek, Joe, and James A. Wilcox. 1991. "The Measurement and Determinants of Single-Family House Prices," *ARUEA Journal*, vol. 19, no. 3, pp. 353-82.

Stephens, William, Ying Li, Vassilis Lekkas, Jesse Abraham, Charles Calhoun, and Thomas Kimner. 1995. "Conventional Mortgage Home Price Index," *Journal of Housing Research*, vol. 6, iss. 3, pp. 389-418.