# Residential real estate price indices as financial soundness indicators: methodological issues

Bradford Case and Susan Wachter[1]

## I. Introduction

The purpose of this conference on real estate indicators and financial stability is "to promote the development of reliable, timely and consistent statistics on real estate prices" in order to support policy initiatives to promote macroeconomic stability. The recent volatility of asset prices and the Asian financial crisis of 1997 have focused attention on the role of asset markets and, in particular, real estate markets in the generation of financial crises and economic instability across nations.[2] With contagion effects driven by fast and large flows of capital, such national crises threaten global financial stability. Hence the need for monitoring devices and policy instruments to respond to the heightened potential for asset market induced global financial instability.

This paper focuses on the potential uses of residential real estate price indices as a tool to monitor asset market instability, and the methodological issues involved in their development. In Section II we examine how real estate price indices can serve as a monitoring device to help minimise financial instability. Section III reviews the methodological issues in the development of residential price indices, and Section IV provides a discussion of the availability of data in the United States to support the development of such indices. Section V discusses what we learn from the price trends revealed in the indices. Section VI concludes.

## II. The use of price indices to monitor asset markets and promote financial stability

Although there are many possible empirical methods and data sources for estimating real estate price indices, not all of them can be expected to play an effective role in promoting financial stability. Before discussing the methodological and data considerations involved in developing a price index, we must consider the function of a properly constructed real estate price index in monitoring asset markets and promoting financial stability.

Fundamentally, the price of any property is equal to the present discounted value of all future services (ie housing) that will be provided by that property while it is owned by its current owner, plus the present discounted value of the price at which the owner will be able to sell the property in the future. In general, we would expect the value of housing services to change only gradually, but the future market price could change more rapidly. To take this a step further, the "market price" of any property at a given time can be defined as the highest price at which the owner would be able to find exactly one willing buyer for that property at that time. This market-clearing price, however, may fluctuate sharply over time: because of changes in the availability of particular types of housing services, because of changes in the cost of financing housing purchases, or because of changes in expectations about future increases in the market-clearing prices among other market participants.

Indeed, a market "bubble" can be thought of as an increase in the market-clearing prices that is justified only by expectations that those price increases will continue into the indefinite future, and not by current or expected changes in the value of housing services or the cost of financing. Although all types of financial instability can be disruptive, it is important to distinguish these market bubbles from

---

[1] The views expressed are those of the authors and do not represent official views of the Board of Governors of the Federal Reserve System or its staff.

[2] See Mera and Renaud (2000).

what may be termed "fundamentally supported" fluctuations in asset values. Both sources of boom-bust cycles are cause for concern, but they may call for sharply different policy responses.

Property markets and real estate prices are inherently subject to booms and busts. One reason for this is construction lags: if a surge in demand pushes the price of existing property above its replacement cost, then developers have an incentive to build more properties. But new properties may take years to complete, and until the new supply is forthcoming, market-clearing prices will remain high. In the presence of construction lags, then, price increases efficiently reflect the current scarcity of housing services. Nonetheless, this increase in market-clearing prices will tend to be followed by a drop once the new supply is forthcoming. This cyclicality in asset values means that lending at any given loan-to-value (LTV) ratio during the price boom - when the demand for construction financing is strongest - may well produce a portfolio of loans with higher than anticipated LTVs when asset values drop after supply responds.

A second reason for cyclicality in asset values is the absence of short selling in real estate markets. Myopic buyers tend to extrapolate price increases into the future, even when sustained price increases are not justified by market fundamentals. In an efficient market, such price rises would be countered by non-myopic investors selling short (that is, selling something for future delivery that they do not currently own, in the hope that they will be able to buy it more cheaply later). But, due to the underlying heterogeneity of properties, there are no organised futures or options markets for individual property sales. In markets with no short sellers, prices will be driven by myopic buyers so long as the upward trend continues. Moreover, as Herring and Wachter (1999, 2002) show, in an economy in which real estate prices have risen over a long period of time with no declines, buyers typically underestimate the likelihood of an eventual downturn. That is, investors are prone to "disaster myopia", the tendency over time to underestimate the probability of low-frequency shocks.[3]

Real estate markets are made more vulnerable to fluctuations because of the role played by the banking system. As Herring and Wachter (1999) show, increases in the price of real estate raise the economic value of bank capital to the extent that banks own real estate; thus banks increase their exposure to real estate when prices are rising. Higher prices also lift the value of banks' own property holdings and hence their capital, which encourages them to relax their lending standards. If prices fall, this process goes viciously into reverse, and a credit crunch can amplify the impact of falling prices.

Moral hazard may also contribute to a bank's supply of capital to real estate, exacerbating booms and busts. To the extent that bank managers' salaries and bonuses are based on reported short-term profits without adjustment for reserves against shocks, the line officers who are in the best position to assess such dangers will be rewarded for disregarding them (Pavlov and Wachter (2004)). Moreover, Pavlov and Wachter (2003) show that, due to competitive pressures in the banking industry, all managers will be pushed to underprice the risk of real estate loans, and, additionally, bank shareholders themselves will incentivise such behaviour.

In addition to problems of moral hazard, poor information and inadequate analysis of real estate risk contribute to the vulnerability of the banking system. Banks and individual managers, besides being poorly incentivised, have little data on which to base careful analysis of future real estate prices. The property value appraisal process is based on observing the prices of comparable properties to estimate the market value of properties (and therefore LTV ratios). While lending decisions would ideally be made on the basis of long-term expectations about the market value of the property throughout the life of the loan, the observed transaction prices of comparable properties are market-clearing prices, subject to bubbles and other sources of short-term fluctuation. Moreover, prices of comparables cannot be used for appraisal purposes until after the transaction is closed, which means that price indices based on appraisals generally lag actual movements in real estate prices.

Real estate price indices can serve in two ways to reduce boom-bust cyclicality in asset value markets, and the attendant cyclicality in the banking system. First, to counter the tendency for banks and appraisers to underestimate LTV ratios by basing them on short-term real estate price booms (whether induced by bubbles or not), indices of current market-clearing prices can be compared to measurements of what might be called long-term property values. "Long-term value", for example, might be thought of as the (relatively stable) value of housing services, plus an average over the range

[3] See Tversky and Kahneman (1982) and Guttentag and Herring (1986).

of future non-bubble market values (all discounted to present value). One advantage of estimating such long-term value is that it could potentially prevent banks from financing property transactions or construction based on unrealistic expectations about future market prices. A major concern, however, is that it is far more difficult to construct indices of long-term value than of current market value. This has consequences for the availability of credit. Transactions occur only at the current market-clearing price: for example, if governments mandate that sales cannot occur at any price above (or below) the long-term value, then sellers (or buyers) will generally refuse to sell (buy) if the current market-clearing price differs from the long-term value.[4] In any case, the tendency for banks to increase exposure to real estate by liberalising LTV ratios during real estate booms can be countered.

Second, extreme volatility in the price index, or extreme differences between the index of current market-clearing prices and "long-term value", can function as a warning that a market bubble has occurred, perhaps requiring a different public policy response. While a discussion of the feasibility of developing such methodologies is a subject for another paper, it may be useful to offer some suggestions on how these analyses could be implemented. For this to work, for example, methodologies could be developed to estimate expected volatility, or the extent of deviations from long-term equilibrium values. Estimation methodologies could be based on housing cycles or on ratios that are derived from such models. Additionally, under simplifying assumptions, ratios of prices to rents could be used to uncover prevailing price change expectations (given real interest rates), which can be compared to model-generated expected price changes. Generally, actual price realisations could be compared to model-specified price outcomes through simulation based on assumptions on supply and demand functional forms. While there are many possible housing market models and the specifications would vary with the underlying characteristics of the economy, all empirical models that are designed to track current asset price realisations as compared to longer-run outcomes must first identify and measure the current market asset price of housing. As the following details, this is not a small task, conceptually or practically.

Finally, it is worth noting that moral hazard and scale economies suggest that the development of real estate price indices is an appropriate exercise of the government's regulatory function. Banks cannot be relied on to construct market-wide price indices, both because they do not individually have adequate data and because their incentive structures may oppose the collection of reliable market data. Furthermore, technology and data requirements mean that there are likely to be strong economies of scale in the development and maintenance of price indices, which suggests the value of centralised price index estimation. While there is certainly a place for private sector estimation of real estate price indices, the goal of financial stability may well best be served by the development of appropriate price indices at the central government level.

## III. Methods used to construct residential real estate price indices

As noted above, there are many possible empirical methods and data sources for estimating real estate price indices, and selecting the most appropriate method and data must depend in large part on the function to be served: monitoring asset market instability and promoting financial stability. Each methodology is usually best suited to a certain type of application. In this section we discuss the different empirical methods available, evaluating the extent to which each method can be expected to further the goal of financial stability.

Four methods are commonly used to compile residential real estate price indices. The most straightforward is simply the *average* or *median price*[5] during each time period. For example, in the United States, the National Association of Realtors publishes an index giving the median price of existing single-family residential properties that transacted in each quarter for each metropolitan

[4] Some governments currently attempt to embody long-term market value in price indices constructed using judgment of local appraisers and assessors. While intuitively interesting from a policy perspective, it is difficult to judge how well such procedures have worked. Moreover, appraisers in the United States and Royal Chartered Surveyors are required to estimate the current market-clearing price, rather then long-term value.

[5] Our discussion focuses on price (at transaction), but indices may also be compiled on the basis of value (at or between transactions). Below we discuss the choice between using transaction prices and values.

statistical area (MSA) in the United States The median price is generally preferred to the average because the distribution of prices is sharply skewed, so that fluctuations in sales volume among expensive properties would have a strong effect on average selling price but a muted effect on median selling price; for some applications, however, average price might be preferable. The data requirements for this method are minimal: simply the prices at which all (or a representative sample of) properties transacted during the time period.

The major problem with this method, however, is quite substantial: it fails to control at all for changes in the quality of the properties whose prices were observed in each period. Quality, of course, tends to improve over time as new properties are constructed, older properties are demolished, and existing properties are renovated; because of this, an average- or median-price index tends to substantially overstate the increase in price for a constant-quality property, or for any existing (and depreciating) property. Moreover, the sample of properties that transacts in a given time period is not constant over time; because of this, an average- or median-price index tends to overstate price increases when all that is happening is that relatively expensive properties are overrepresented among transactions, and understate price increases when relatively inexpensive properties are overrepresented.

A second reasonably straightforward technique to track property prices, the *representative-property* method, is implemented by defining a representative property and then recording in each period the price (or value) of a property conforming to the specified characteristics. The shelter component of the US consumer price index (CPI) essentially proceeds in this way, as do some proprietary indices such as the National Real Estate Index published by Global Real Analytics. The only data item that is actually collected is the price of the representative property in each time period. In order to implement this method, the data collector must observe all of the characteristics used in defining the representative property in order to select one conforming to the definition. The major problem with this method is that data points may not be fully comparable across markets or over time, because of differences among data collectors in subjectively interpreting the definition of the representative property and applying that definition to choose a representative property. A second problem is unmeasured quality change: specifically, quality improvements that are not captured by the definition of the representative property. For example, if a property is defined with respect to location, lot size, living space, and number of rooms but not with respect to major amenities such as central air conditioning, then any increase in the prevalence of those amenities will show up improperly as an increase in the price index rather than properly as an improvement in quality. Finally, because the method focuses on the price of only one property (the representative), it does not take advantage of information contained in the prices of all other properties; in fact, in extreme cases it may not reveal the movements in the general price level if, for whatever reason, the representative property's price does not respond in the same way.

In order to avoid the problems inherent in the average-/median-price and representative-property methods, economists estimate price indices using *hedonic-price* models. These models postulate that the transaction price of any given property is a function of the time period in which it transacted as well as its hedonic characteristics - that is, the physical features of the house or lot, and the features of its location and neighbourhood, that affect the price at which it transacts. If we know the hedonic function, then regression analysis can be used to estimate the parameters of this function. For example, a common hedonic-price function is

$P_{it} = \alpha X_i^{\beta_1} e^{\beta_2 Y_i + \gamma_1 T_{i1} + \gamma_2 T_{i2} + \ldots + \gamma_n T_{in}}$ or, in logs,

$$\ln P_{it} = \alpha + \beta_1 \ln X_i + \beta_2 Y_i + \gamma_i T_{i1} + \gamma_2 T_{i2} + \ldots + \gamma_n T_{in}$$

where $P_{it}$ is the transaction price of property $i$ during time period $t$; $X_i$ and $Y_i$ are hedonic attributes of the property (with $X$ measured continuously - say, square feet of living space - and $Y$ measured discretely - say, presence of central air conditioning); $T_{i\tau}$ are dummy variables indicating whether the transaction took place during time period $\tau$; and $\alpha$, $\beta_j$ and $\gamma_\tau$ are the parameters to be estimated. In particular, the series of parameters $\gamma_\tau$ is the price index.

The hedonic-price method offers several advantages over the average-/median-price and representative-property methods. First, and most importantly, the hedonic-price method controls for quality change: specifically, if there has been any change in the attributes measured $X_i$ and $Y_i$ - either because the quality of individual properties has changed or because different-quality properties are more likely to transact - then this quality change will be reflected in the hedonic measures rather than in the parameters (including the price index). This is a great advantage over the use of the median and average price. Compared to the representative-property method, the hedonic-price method does not

require subjectively interpreting the definition of the representative property or applying that definition to choose a representative property, nor does it fail to make use of data from other properties.

On the other hand, the method does have some disadvantages as well. The data requirements are much more onerous than for the average-/median-price method,[6] as the analyst should have data on all of the hedonic attributes of the property, as well as its price, at the time of the sale. This method potentially shares the problem of unmeasured quality change; if the hedonic measures do not capture amenities that improved over time, then any increase in the prevalence of those amenities will show up improperly as an increase in the price index rather than properly as an improvement in quality.

Another disadvantage is the problem of determining the correct model specification. The hedonic-price function must be specified correctly - that is, the analyst must use the correct "functional form" and include all relevant hedonic characteristics (ie must not have any omitted variables) in order to generate unbiased estimates of the price index.[7] Moreover, the parameters on the hedonic-price attributes ($\beta_j$, called the "implicit market prices" of the attributes) must not have changed over time, or if they have, then that must be incorporated into the functional form. Any violation of these conditions - incorrect functional form, omitted variables or changing parameters - *theoretically* will result in biased estimates of the price index. In practice, however, it appears that the hedonic-price method is quite robust to reasonably minor violations of these conditions: for example, it appears that the estimated price index will be fairly close to the true price index as long as several of the most important hedonic attributes (eg number of bathrooms) are included. Thus it seems, in practice, that the major disadvantage associated with the hedonic-price method is the cost of data collection.[8]

The onerous data requirements of the hedonic-price method (as well as of the representative-property method) have encouraged analysts to use a simpler method derived from hedonic-price models, called the *repeat-sales method*. This method takes advantage of the fact that when a given property transacts twice, many or most of the hedonic attributes of that property will not have changed at all between transactions. To the extent that this is true, the analyst need not collect data on the level of each hedonic attribute at the time of either sale; it is enough to know that the attribute has not changed. In these cases, the change in price of the property between transactions can be expressed as a simple function of the time periods elapsed between transactions. The cost and ease of implementation advantages of the repeat-sales method have made it the price index methodology of choice for large-scale applications: for example, price indices for single-family residential properties in several hundred US metropolitan areas (as well as at the national, regional and state levels) are published quarterly by both Freddie Mac and the Office of Federal Housing Enterprise Oversight (OFHEO). Nonetheless, as discussed further below, there are important measurement disadvantages in the use of such indices. Chief among these disadvantages is the need for frequent transactions. The repeat-sales methodology can only be used in markets where properties are transacted frequently and plenty of sales data are available. In western Europe, for example, the repeat-sales methodology is not useful given the small number of housing transactions. Moreover, it should be noted that repeat-sales price indices need to be combined with initial priced hedonic indices to compute comparable price levels across markets.

The repeat-sales model is derived from the hedonic-price model by expressing the ratio of the prices for two transactions of the same property as the ratio of the right-hand-side hedonic functions for those two transactions:

---

[6] But no more onerous than the representative-property method: although only the price of the representative property is actually recorded, information on the full set of hedonic characteristics should be used to define and identify a representative property.

[7] Halvorsen and Pollakowski (1981) addressed the difficulty of selecting the correct functional form for a hedonic price model. Meese and Wallace (1991) proposed a non-parametric method for estimating the implicit market prices of hedonic attributes in order to avoid this problem.

[8] Constant-quality methodologies are ideal for many uses and applications such as attempting to identify a "bubble" in housing markets. In this type of analysis, the pure price signal is what should be identified and analysed in an attempt to see if pricing has become irrationally high. Nonetheless, a financial institution attempting to "mark to market" LTV ratios on a portfolio of mortgages would not want to use a "constant-quality" methodology since improvements in quality on a collateral property are real improvements in value that should be considered.

$$\frac{P_{it}}{P_{it'}} = \frac{\alpha X_i^{\beta_1} e^{\beta_2 Y_i + \gamma_1 T_{i1} + \gamma_2 T_{i2} + \ldots + \gamma_n T_{in}}}{\alpha X_i'^{\beta_1} e^{\beta_2 Y_i' + \gamma_1 T_{i1}' + \gamma_2 T_{i2}' + \ldots + \gamma_n T_{in}'}}$$ or, in logs,

$$\ln \frac{P_{it}}{P_{it'}} = \beta_1 \ln \frac{X_i}{X_i'} + \beta_2 (Y_i - Y_i') + \gamma_1 (T_{i1} - T_{i1}') + \gamma_2 (T_{i2} - T_{i2}') + \ldots + \gamma_n (T_{in} - T_{in}').$$

Because $X_i' = X_i$ and $Y_i' = Y_i$, this can be simplified to

$$\ln \frac{P_{it}}{P_{it'}} = \gamma_1 (T_{i1} - T_{i1}') + \gamma_2 (T_{i2} - T_{i2}') + \ldots + \gamma_n (T_{in} - T_{in}')$$

where $P_{it}'$ is the transaction price at the time of the previous sale; $X_i'$ and $Y_i'$ are the hedonic attributes of the property at the time of the previous sale; $T_{i\tau}'$ are dummy variables indicating whether the previous transaction took place during time period $\tau$; and the series of parameters $\gamma_\tau$ is the price index. Two points are worth noting about the right-hand side of this equation. First, the expressions $T_{in} - T_{in}'$ take the values –1 during the time period of the first transaction, +1 during the time period of the second transaction and 0 otherwise. Second, any property that sold at least twice[9] *in different time periods* can be included in the analysis, but if all transactions of that property occurred during the same time period, then the property must be dropped from the analysis because all terms on the right-hand side will have the value 0.

As noted, the major advantage of the repeat-sales model is that it requires little data collection. This applies so as long as it is known that none of the relevant hedonic characteristics of the property have changed between transactions. However, it is easy to overestimate this advantage in practice, because an analyst must have some reliable way to determine whether, indeed, the property's characteristics have remained constant. This generally means that the practical data requirements of the repeat-sales model are quite similar to those of the hedonic-price model or, alternatively, the potential that an index increase is simply due to quality increases cannot be determined.

In practice, analysts generally assume that hedonic attributes have *not* changed between transactions, and this assumption of course greatly reduces the data collection burden. This assumption, however, is not generally true, and for this reason the unmeasured quality change introduces an unknown positive bias into the estimated price index, thus undermining its use in monitoring unsustainable price increases.

An advantage of the repeat-sales model is that it automatically controls for *all* hedonic characteristics that remained unchanged between transactions, whereas the hedonic-price model controls only for those that are measured. Because of this, the repeat-sales method makes more efficient use of the information contained in repeat transactions of a given unchanged property. There is a major cost associated with this, however: because it uses information only on transactions of those properties that sold at least twice during the study period (and remained unchanged between transactions), the repeat-sales method ignores a very large amount of potential information from transactions of properties that sold just once during the study period (or that changed between transactions). The number of property transactions ignored in this way varies with the length of the study period and the level of market activity, but generally is the great majority of available transactions.[10]

Another disadvantage of the repeat-sales model is the changing-parameters problem discussed above in connection with hedonic-price models: the parameters on the hedonic-price attributes ($\beta_j$, the implicit market prices) must not have changed over time, or if they have, then that must be incorporated into the functional form of the hedonic-price equation. In the standard repeat-sales formulation, however, there is no way to modify the functional form of the equation to incorporate

---

[9] If a property has transacted more than twice during the study period, then each observation (transaction pair) on that property must be weighted to correct for collinearity in the disturbance terms. See Bailey et al (1963) and Palmquist (1982).

[10] Moreover, properties may be more or less likely to transact depending on whether prices are increasing rapidly or slowly (or declining), in which case a repeat-sales price index could potentially be biased. Note that this problem of sample selection bias would also exist for hedonic-price methods, but would be much less serious even than for repeat-sales methods.

changes in implicit market prices; instead, the effect of the changed implicit market prices will be improperly reflected in the price index.[11]

This disadvantage, together with the other shortcomings of the repeat-sales method - failure to use information from properties that transacted just once, (measurable) changes in hedonic attributes, and changes in implicit market price parameters - motivated the development of a hybrid model that combines attributes of both the repeat-sales and the hedonic-price method.[12] The essence of hybrid models is that they "stack" repeat-sales and hedonic models, and then estimate the two models imposing a constraint that estimated price changes over time are equal in both models. In effect, such methods are weighted averages of the hedonic and repeat-sales methods. The evidence suggests that repeatedly sold properties may differ in unobserved ways from other properties, in which case the stacking method induces measurement error. While such indices, unlike hedonic or repeat-sales indices, do use all available information, Case et al (1991) do not find clear efficiency gains from using the hybrid model instead of the hedonic approach.

## IV. Types of price data with which to create residential price indices

In addition to the wide variety of empirical methods available, there are also many different types and sources of data that could be used to construct real estate price indices. As we discuss in this section, however, few of these data sources would support the development of reliable price indices that can be expected to promote the goal of financial stability.

In the United States, there are several sources of data on real estate prices or values, some collected by government agencies and provided to the public free of charge, others collected privately and kept private or offered for sale. The Census Bureau of the US Department of Commerce, for example, provides data on sales price, and median and average prices, on an annual and quarterly basis, for *New Houses Sold* and another set of price indices for *Median Prices of Existing Family Dwellings*.[13] The major limitation of these data series is the overstatement of house price appreciation, because they do not account for the changes in quality that occur over time. The Census also constructs, based on new construction, a *Constant-Quality Pricing Index*, since 1978, although its value is lessened due to its geographical limitations. A set of statistical models relating sales price to selected standard physical characteristics of house units is used to derive the average price for a constant-quality unit. Generally, the geographic distribution of these indices is limited to an aggregate index of the United States and the four major census regions.[14] An issue to be considered when prices are based off new construction is where the new construction occurs. Because new construction is likely to occur on the fringe of urban areas where supply is elastic, such indices may underestimate property price appreciation. For example, price appreciation of newly constructed homes in suburban Rhode Island or Massachusetts would not be comparable to the appreciation rates of condominiums in downtown Boston. New construction methodologies therefore may not pick up the effects of land scarcity in a market, and may tend to underestimate overall market price appreciation for this reason.

A second US government source for house price data, the CPI published by the Bureau of Labor Statistics of the US Department of Labor, is constructed using the representative-property method.[15] The largest part of the CPI's housing series is in the shelter category,[16] which covers rent of primary

---

[11] Shiller (1993) showed that a generalisation of the standard repeat-sales formulation, however, permits the estimation of a separate price index for each hedonic attribute.

[12] See Case and Quigley (1991).

[13] These data are constructed by the National Association of Realtors (NAR), who also release a quarterly, quality-unadjusted series for a panel of large MSAs based on median prices from the local Board of Realtor Multiple Listing Service.

[14] In addition, the decennial census data record house prices and rents, and publish median values for MSAs and even smaller jurisdictions.

[15] See http://www.bls.gov/cpi/cpifact6.htm for more information.

[16] Other parts include the price of household furnishings, appliances, utility services, etc.

residence[17] and owner's equivalent rent (far and away the most heavily weighted item in the overall series). Owner-equivalent rent is constructed from data provided by homeowners themselves. Homeowners are asked what their unit would rent for.[18] This methodology appropriately calculates changes in owner user costs and, by design, it does not measure changes in house prices or values. As discussed above, change in the owner-equivalent rental component of the CPI can be compared to value change as an indicator of asset price inflation relative to changes in the price of the underlying stream of housing services, but it cannot be used to measure house price inflation.

A third source of price data is mortgage transactions, which are used for repeat-sales price indices provided by OFHEO. As federal regulator of the two government-sponsored enterprises (Freddie Mac and Fannie Mae), OFHEO has access to the nation's largest database of mortgage transactions, over 23 million repeat transactions on conforming, conventional single-family loans insured by the GSEs. Both OFHEO and Freddie Mac make quarterly series, organised by census division, state, MSA, or national, available on their website, free to the public. The indices are easily downloadable in Excel or text format, generally two months after quarter-end. The national, census division, and state series are available back to 1975, but the MSA series have different starting points because an MSA series is only published if at least 1,000 observable transaction pairs exist in the area. One advantage of these data is the high frequency, but this also leads to frequent revisions. Each quarter, recent mortgage transaction data from the GSEs are combined with past data and calculations are then performed on this updated dataset. The index is created by looking at all properties which have been sold at least twice and comparing the two sale prices using a modified Case-Shiller method.[19]

A disadvantage of the OFHEO and Freddie Mac series, besides those discussed above that are generic to repeat-sales price indices, is their limited geographical coverage. Private analysts, such as Case and Shiller (1989, 1990), present basic results for an additional but still limited number of locations. The private firm Fiserv CSW (formerly Case-Shiller-Weiss) and a collaboration of the research departments of Fannie Mae and Freddie Mac have produced such indices for a wide range of MSAs and smaller areas; however, the small area indices are proprietary and not readily available for research purposes.

Most of the discussion in this paper has been phrased in terms of a price index based on property transactions, but that is not the only type of data that can be used (or that is commonly used) to compile residential real estate price indices. The advantages of using actual sales prices from property transactions are, first, that sales prices (from arm's length, non-coerced transactions) represent the most reliable indicator of the actual market value of any given property; and, second, that data on sales prices may already be readily available if they are collected for the administration of real property taxes, transfer taxes, deed recording fees, or other public purposes. The disadvantages of using sales prices from property transactions are, first, that during a given study period only a fraction (generally a small fraction) of all properties will have transacted even once; and, second, that if some properties are more or less likely to transact depending on whether prices are increasing rapidly or slowly (or declining), then the use of transaction prices may introduce sample selection bias into the estimation of the price index. These disadvantages appear to be minor compared to the quality advantage of data from actual market transactions.[20]

It is also possible, however, to compile residential real estate *value* indices from observations on what is believed to be the underlying market value of a given property, as opposed to the price observed (only) when that property transacts. Perhaps the most straightforward source is estimates of the market value of each property that are recorded for the purposes of assessing real property taxes,

---

[17] Shelter also includes lodging away from home, housing at school, excluding board, other lodging away from home including hotels and motels, and tenant's and household insurance.

[18] From 1987 to 1998, CPI data collectors gathered information from the owners to calculate an appropriate initial, implicit rent. Changes for similar (based on structure and attributes) renter-occupied units were then applied to the initial value to calculate changes in owner-occupied implicit rents. Since 1998, the rent index of the survey has simply been reweighted to rents on the CPI Housing Survey.

[19] http://www.ofheo.gov/Media/Archive/house/hpi_tech.pdf.

[20] For owner-occupied multifamily rental properties, price indices that are based on transactions may reflect variation in liquidity over the business cycle, which affects the ease with which owners are able to sell properties. Fisher et al (2003) propose a constant-liquidity price index method, and find that movements in the constant-liquidity index tend to lead movements in a transaction-based index.

which are imposed almost universally in the United States. Real property tax assessment records are readily available in any jurisdiction that imposes the real property tax, and they are established regularly, generally every year. Unfortunately, real property tax assessment records are generally considered to be of very poor quality. Even if they are updated annually, the updating process may bear little relation to changes in the market price level, for several reasons. First, for simplicity most jurisdictions adjust the assessed values of all properties within the jurisdiction by the same factor, regardless of whether prices in parts of that jurisdiction have increased more or less rapidly than the average. Second, the adjustment factor is set through a political process that need not reflect actual market fluctuations. Third, assessed values for individual properties may be set closer to market prices only when those properties transact; indeed, in some jurisdictions (such as California) assessed values may be explicitly prevented from adjusting to the same extent as market prices. Finally, homeowners are much more likely to challenge the estimated values on which their property tax assessments are based when those values have increased sharply, so property tax assessment records generally tend to understate the actual pace of property value increases. For these reasons property tax assessments are rarely, if ever, used as residential real estate value indices in themselves.

A much more commonly used source of market values is records from private appraisals, which are generally conducted in connection with mortgage transactions - whether purchase-money mortgages upon a property transaction, or mortgage refinancings. Indeed, in the United States the indices published by Freddie Mac and OFHEO both include appraised values from records of refinanced mortgages purchased by Freddie Mac (and, for the OFHEO index, Fannie Mae). The quality of private appraisals is probably much higher than the quality of property tax assessment records, but appraisals may still differ sharply from underlying market values because of the subjectivity of the appraisal process, particularly if the subject property did not transact at the same time and there were few transactions of closely comparable properties during the same time period. Moreover, appraisal-based price indices may suffer from sample selection bias, especially since homeowners may be more or less likely to refinance their mortgages if property values are increasing rapidly.[21] For these reasons, economists have found that price indices based on appraisal records tend to be smoother than price indices based on transaction records and tend to misrepresent the times at which market prices reach their peaks or troughs. Also for these reasons, in the United States both Freddie Mac and OFHEO are considering deleting appraisals that were conducted in connection with mortgage refinancings from the computation of their price indices.

Another source of information on property values is records on listing prices of properties offered for sale: for example, various local multiple listing service (MLS) databases in the United States have been used to construct value indices. The advantages of these data are that (1) listing prices are established with the assistance of real estate agents, who may have especially good judgment regarding the value that would be assigned to each property in a well functioning market, and (2) the number of properties listed for sale during any time period is even greater than the number of property transactions. The disadvantages of these data, however, are closely related to the advantages. First, listing prices may differ sharply from underlying market values, partly because neither real estate agents nor homeowners may be good judges of market value and partly because they may in fact have incentives not to equate the listing price with the market value. Second, properties with particularly low listing prices relative to market value can be expected to transact quickly, while properties with particularly high listing prices relative to market value can be expected to remain on the market for a long time and perhaps never transact. For these reasons listing prices are generally not considered a reliable source of market value data.

Nonetheless, the underlying data collected on assessments and listing prices have themselves been used in estimation of hedonic indices (Clapp and Giacotto (1992)). Moreover, the underlying data on sales transactions, including prices, date of sale, and housing attributes, collected by the MLS and by municipalities, are potentially valuable for the construction of price indices. Appraisal and assessor

[21] In addition, some property appraisals may be motivated not by mortgage transactions but simply by the observation that the pace of market price increases seems to have changed significantly. This is much less common among (single-family) residential properties than among commercial properties (including multifamily residential properties), but should be recognised because price indices based on appraisals that are motivated in part by sharp increases, or declines, in the general property price level can be expected to suffer from sample selection bias.

agencies are moving towards using these data for statistical-model based price estimation. Assessors are incorporating hedonic methodologies in computer-assisted mass appraisal (CAMA) and appraisers are using automated valuation methods (AVMs) for desk review appraisals and for mortgage underwriting. Thus lenders and municipal authorities are increasingly making use of statistical methods to estimate the market value of homes; these technologies also have the potential to generate standardised[22] hedonic local area residential price indices.

A final source of market values is simply a survey of homeowners requesting that they assess the value of their own properties. The American Housing Survey, for example, records owner-assessed property values, and these values have been used to construct value indices.[23] In principle, owner-assessments can provide a useful source of market value information, as homeowners (1) are particularly knowledgeable about the condition and amenities (structural and locational) of their properties, and (2) often observe market prices of comparable neighbouring properties. However, property owners are not necessarily good judges of the value that would be assigned to their properties by a well functioning market. Indeed, economists have found that homeowners tend to overestimate the market values of their properties, and tend furthermore to underestimate the rate of increase in the market values of their properties (Kiel and Zabel (1999)). For this reason owner assessments are generally not considered a reliable source of market value data with which to construct property value indices. Nonetheless, these data have been used by researchers to construct hedonic indices for the United States. For example, Malpezzi et al (1980) used AHS data from the 1970s to construct constant-quality indices for a sample of MSAs. This work was subsequently updated and expanded by Thibodeau (1992, 1995).

## V. What stories do the US residential real estate price indices tell?

The most important story told by residential price indices about the US residential real estate market is that, in the long run, housing price increases in the United States have tracked increases in the overall price level quite closely. Graph 1 below shows the overall CPI, CPI Rent, CPI Owner-Equivalent Rent and Census Constant-Quality price indices from 1979 to 2002. The growth rates of these indices over this roughly 20-year period were similar; however, the close relationship between housing prices and overall prices often does not hold over short time periods. For example, in Graph 2, year-over-year growth rates in the CPI, CPI Housing and Census Constant-Quality indices are shown from 1997 to 2003. In six of the seven years since 1997, the appreciation in the CPI Housing index has exceeded the growth in the overall CPI index. However, using the growth of the Census Constant-Quality index as a measure, the housing price growth rate was significantly higher than the growth of the overall CPI and CPI Housing indices only in 2002.

In order to determine if there is currently a “bubble” in US residential real estate, it is important to look not only at housing price increases, but also at rent increases. If house prices are appreciating rapidly, this does not necessarily imply that a “bubble” exists if rent prices are increasing just as rapidly, since consumers are rationally pricing increasing rents into owner-occupied housing units.

The data in Graphs 1 and 2 do not demonstrate the existence of a bubble in US residential real estate markets. Over the past seven years shown in Graph 2, rents and constant-quality appreciation have been very similar, and in most years appreciation in rents (CPI indices) has actually been higher than growth in constant-quality home prices. While the ratio of CPI Rent index to the housing price, using the Constant-Quality index, does not show any decline, some private data collected on rents do imply declining rents over 2001 and 2002 that, when coupled with increasing constant-quality house prices, could lend some strength to arguments that a “bubble” does exist. Private data compiled by REIS, RERI and others show rents declining over 2001 and 2002. These indices include only effective rents on newly leased properties, and do not consider rental increases on properties which are already

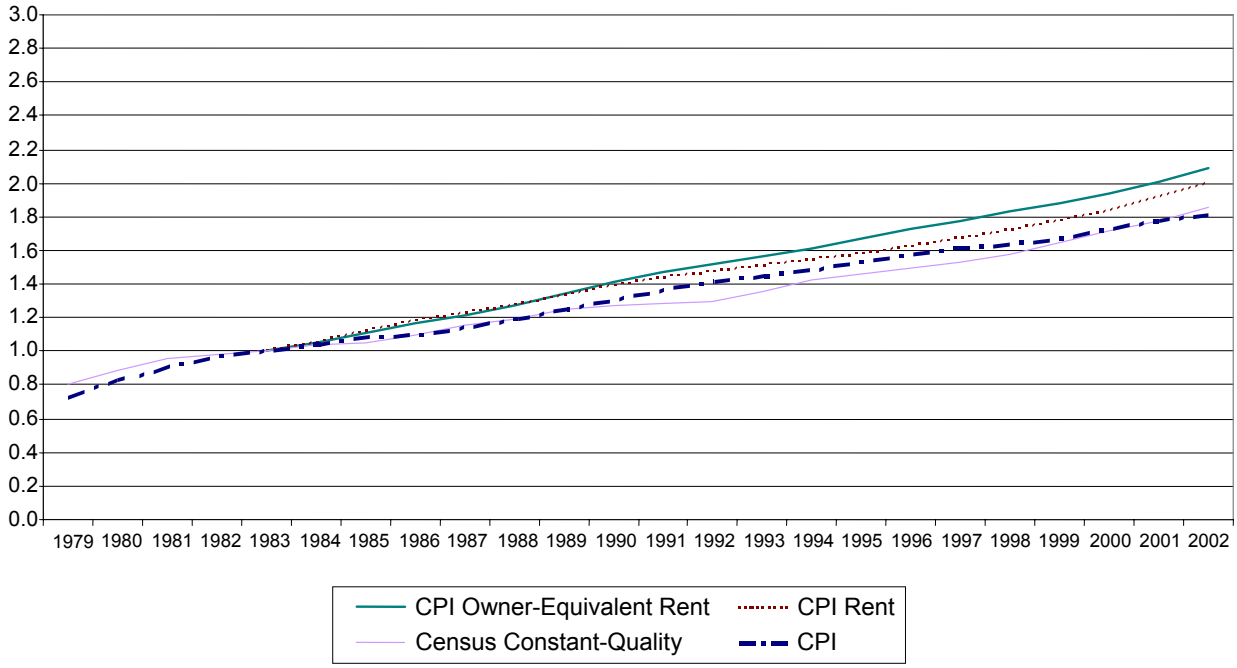
[22] Some municipalities collect information on numerous housing attributes, others on few. However, the use of geographical information, which is available for all municipalities, can substitute for an inclusive list of attributable variables.

[23] To supplement the decennial census, the Commerce Department releases the American Housing Survey, started as an annual survey in 1973 and changed to a biannual one in the early 1980s. Currently, the AHS covers about 50,000 housing units throughout the United States.

leased. These data therefore may more accurately reflect the current state of rental markets, rather than the "smoothed" CPI indices that include rents and escalations on existing leases.

Graph 1

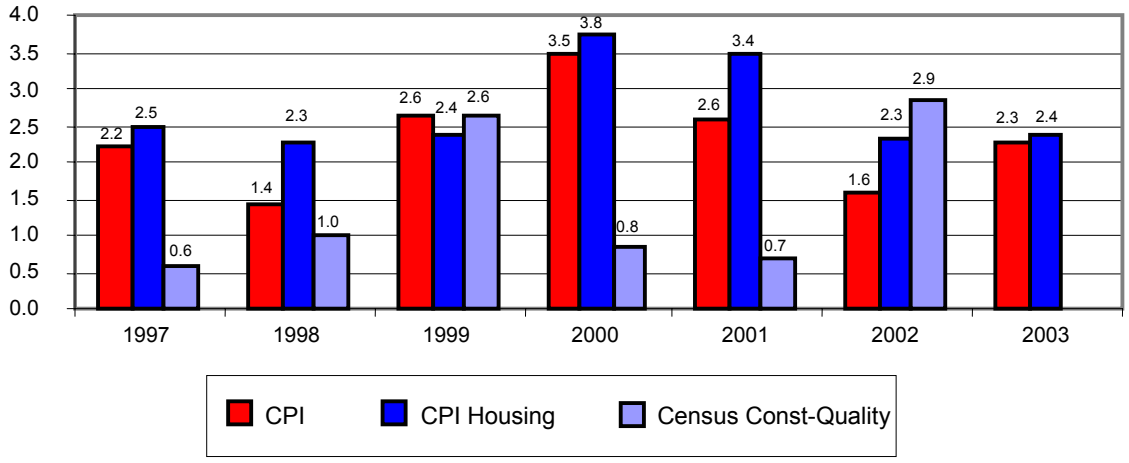
**CPI vs Census Constant-Quality, CPI Rent and CPI Owner-Equivalent Rent - 1979-2003 (1983 = 1.0)**




Graph 2

**Annual growth in overall CPI, CPI Housing and Census Constant-Quality Indices, 1997-2003**

September-September (in per cent)

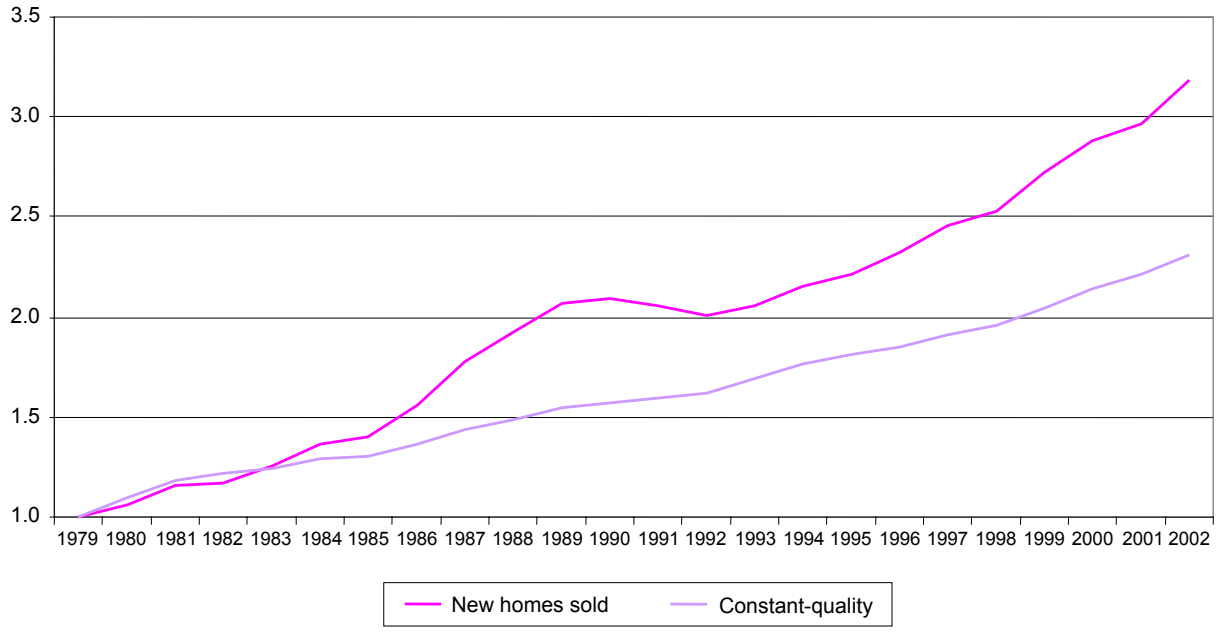


It should also be remembered that the recent accelerating growth in constant-quality prices shown in Graph 2 above is within a range of increases that would be predicted given the tremendous declines in mortgage rates experienced over the past 10 years. While some observers consider the effects of declining interest rates on housing prices a "bubble", it is important not to confuse a "bubble" with a commonly experienced cycle. Bubbles usually refer to irrational asset pricing, but consumers have been completely rational to bid up home prices as interest rates have declined. However, this is not to say that housing prices will not experience some weakness as the cycle turns, and consumers bid lower amounts due to increasing interest rates.

While "headline" price indices, such as means and medians, have shown rapid growth in recent years, it is important to remember that these numbers are influenced by increasing quality of housing, and are not representative of pure price inflation. In Graph 3, a price index for new homes sold is compared to a constant-quality price index. The much greater increase in new home prices when compared to the increase in constant-quality prices shows that Americans are increasingly demanding much better quality in their housing, and that this demand is driving overall housing transaction prices higher. However, one must remember that this quality-related appreciation is not a "bubble", since consumers are paying more for a better product, a completely rational economic behaviour.

Graph 3

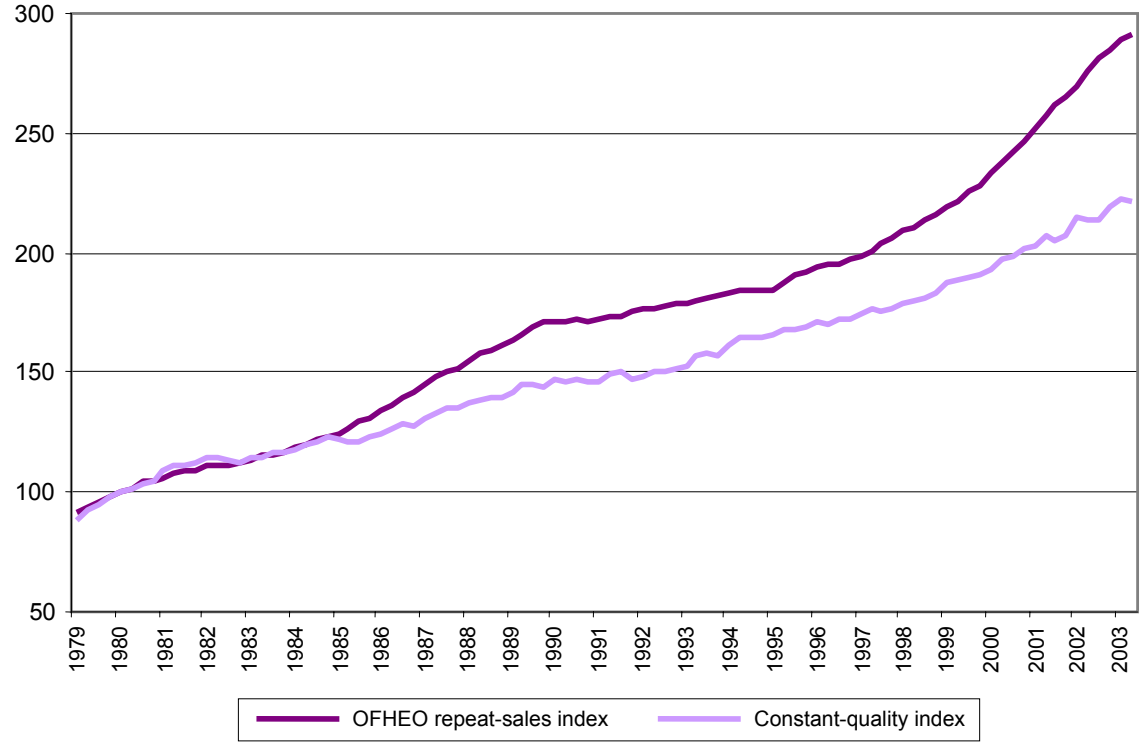
**Price indices for new homes sold vs constant-quality price index**
**1979-2003**




Many people feel that repeat-sales indices control for changes in housing quality, but, in reality, this is not the case. The quality of a single house is not static between transactions, since owners may renovate, expand, or make other quality improvements to the property. The data shown in Graph 4 bear out this hypothesis. In the graph, the OFHEO repeat-sales index is compared to a constant-quality index. Since 1985, the OFHEO index has increased much more rapidly than the constant-quality index, showing that repeat-sales indices do not fully control for changes in housing quality since owners may improve quality between transactions. While this does show real positive investment in the nation's housing stock, this is not to say that this investment in quality will continue indefinitely. If interest rates increase, demand for real estate may decline, and the current investment in real estate quality may prove excessive.

Graph 4

**OFHEO repeat-sales index vs constant quality index, 1979-2003**




One important caveat to the above analysis is that all the indices used were national, and while they do not seem to imply the existence of a national residential real estate bubble, it does not follow from this that bubbles do not exist in any individual regional markets. Real estate markets are regional in nature, and anyone who wishes to analyse the state of the market should rely more heavily on regional price indices of interest rather than aggregated national indices. Individual markets or regions can have vastly different current situations and historical experiences with real estate pricing and appreciation than does the nation as a whole.

## VI. Conclusion

The organisers of this conference have recognised the fundamental connection between real estate markets and financial stability, and therefore the need for prudential supervision and monitoring of real estate markets. Because banks are exposed to cyclicality in real estate markets, and because banks' incentive structures may lead them to exacerbate boom-bust cycles in real estate markets, fluctuations in real estate prices have the potential to strain financial stability and even to jeopardise entire financial systems. In countries in which banks play a dominant role - such as Japan, where banks hold some four fifths of total assets - the consequences for the real economy can be severe.

In particular, for several reasons, banks are liable to increase their origination of real estate loans at the same time that short-term, market-clearing asset prices are at their peaks. As asset prices revert to their longer-term values, however, the result is that banks hold portfolios of loans with higher LTV ratios than anticipated. To counter this tendency - whether it is associated with market bubbles or simply "fundamentally supported" price fluctuations - it is necessary to continually monitor real estate markets, in particular to challenge weakening of underwriting standards when short-term asset prices are rising. This task requires the development of reliable real estate price indices.

There is a wide variety of empirical methods and data sources that could be used to construct real estate price indices; as we point out in this paper, however, not all can be expected to support the goal of financial stability. One straightforward method, for example, simply reports the average or median price of houses transacting during each time period. This method, however, fails to control at all for quality changes or for changes in the mix of transacting properties; thus it presents a picture of asset price movements that is both biased upwards (because quality increases over time) and unreliable (because the mix of transacting properties may change during different parts of the market cycle). A second straightforward method, reporting the price of a representative property, is not well suited for measuring residential property asset value, since such properties transact infrequently.

The hedonic-price method offers a way of avoiding the quality change, comparability and narrowness problems associated with the first two methods; unfortunately, the data required to estimate a hedonic-price model make this method relatively expensive to implement. Because of this, perhaps the most reliable price index method in wide use in the United States, for the nation as a whole as well as for the states, is the repeat-sales method, which requires only two pieces of data (transaction price and date) along with the troublesome assumption that no relevant characteristics of the property changed between transactions. Hybrid models offer the potential to improve on repeat-sales methods where additional data are available, but have not yet been shown to be appreciably superior to repeat-sales methods.

Several data sources could be used to estimate real estate price indices, but many of these are unsuitable for the purposes of monitoring asset markets and promoting financial stability. Owner assessments of property value, property listing prices, and property tax appraisals all suffer from severe problems of bias and unreliability. The best source of data is records of property transaction prices and dates. In the United States, these records are commonly collected and made public in local real property tax assessment systems, many of which also contain records of hedonic property characteristics, thus offering the potential of hedonic-based residential price indices for local areas.

## References


Bailey, Martin, Richard Muth and Hugh Nourse (1963): "A regression index method for real estate price index construction", *Journal of the American Statistical Association*, vol 58, pp 933-42, December.

Case, Bradford, Henry Pollakowski and Susan Wachter (1991): "On choosing among house price index methodologies", *American Real Estate and Urban Economics Association Journal*, vol 19.3, pp 287-307, Fall.

Case, Bradford and John M Quigley (1991): "The dynamics of real estate prices", *Review of Economics and Statistics*, 73(1), pp 50-8.

Case, Karl and Robert Shiller (1989): "The efficiency of the market for single-family homes", *American Economic Review*, vol 79, issue 1, pp 125-37.

——— (1990): "Forecasting price and excess returns in the housing market", *American Real Estate and Urban Economics Association Journal*, 18, pp 253-73.

Clapp, John and Carmelo Giacotto (1992): "Estimating price indices for residential property: a comparison of repeat sales and assessed values methods", *Journal of the American Statistical Association*, 87(418), pp 300-6, June.

Fisher, Jeffrey, Dean Gatzlaff, David Geltner and Donald Haurin (2003): "Controlling for the impact of variable liquidity in commercial real estate price indices", *Real Estate Economics*, 31(2), pp 269-303.

Guttentag, Jack and Richard Herring (1986): "Disaster myopia in international banking", *Princeton University Essays in International Finance*, no 164, September.

Halvorsen, Robert and Henry O Pollakowski (1981): "Choice of functional form for hedonic price equations", *Journal of Urban Economics*, 10, pp 37-49.

Herring, Richard and Susan Wachter (1999): "Real estate booms and banking busts - an international perspective", Group of Thirty, *Occasional Papers* 58.

——— (2002): *Real estate bubbles, asset price bubbles: implications for monetary, regulatory, and international policies*, George Kaufman (ed), MIT Press.

Kiel, Katherine and Jeffery Zabel (1999): "The accuracy of owner-provided house values: the 1978-1991 American housing survey", *Real Estate Economics*, 27(2), pp 263-98.

Malpezzi, Stephen, Larry Ozanne and Thomas Thibodeau (1980): *Characteristic prices of housing in 59 SMSAs*, Washington DC, Urban Institute Press.

Meese, Richard and Nancy Wallace (1991): "Nonparametric estimation of dynamic hedonic price models and the construction of residential housing price indices", *AREUEA Journal*, 19(3), pp 308-32.

Mera, Koichi and Bertrand Renaud (2000): *Asia's financial crisis and the role of real estate*, M E Sharpe, Armonk.

Palmquist, Raymond (1982): "Measuring environmental effects on property values without hedonic regressions", *Journal of Urban Economics,* 11, pp 333-47.

Pavlov, Andre and Susan Wachter (2003): "The anatomy of non-recourse lending", *Wharton Real Estate Center Working Paper*, July.

——— (2004): "Robbing the bank: non-recourse lending and asset prices", *Journal of Real Estate Finance and Economics*, February.

Shiller, Robert (1993): "Measuring asset values for cash settlement in derivative markets: hedonic repeated measures indices and perpetual futures", *Journal of Finance,* 68(3), pp 911-31, July.

Thibodeau, Thomas (1992): *Residential real estate prices: 1974-1983*, Blackstone, Mt Pleasant, Mich.

——— (1995): "House price indices from the 1984-1992 MSA American Housing Surveys", *Journal of Housing Research*, 6(3), pp 439-79.

Thompson, Michelle, Kevin Gillen and Susan Wachter (2003): "GIS for real estate valuation", *Wharton Real Estate Center Working Paper*.

Tversky, Amos and Daniel Kahneman (1982): "Availability: a heuristic for judging frequency and probability", in D Kahneman, P Slovic and A Tversky (eds), *Judgement under uncertainty: heuristics and biases*, pp 163-78.