Conventional Mortgage Home Price Index

William Stephens, Ying Li, Vassilis Lekkas, Jesse Abraham, Charles Calhoun, and Thomas Kimner*

Abstract

This article presents the background work for the Conventional Mortgage Home Price Index, a new set of indices published quarterly by Fannie Mae and Freddie Mac. This version of the indices uses a sample of repeat mortgage transactions for single-family properties in a national database of loans purchased or securitized by the two agencies between 1975 and 1992. The indices are estimated by the repeat-sales method, which can produce "constant-quality" indices even without detailed property characteristics because it uses information on the same properties at two points in time.

Indices based on agency repeat transactions are subject to statistical variation caused by renovation bias, transaction bias, refinancing bias, and revision volatility, but the agency index appears to overcome these problems. Comparisons with other indices indicate that the Conventional Mortgage Home Price Index is a cost-effective way to achieve both wide coverage and geographic disaggregation in measuring house price changes.

Keywords: constant-quality price; Conventional Mortgage Home Price Index; Fannie Mae; Freddie Mac; repeat-sales index

Introduction

This article presents the background work for the Conventional Mortgage Home Price Index (CMHPI), a new set of home price indices published quarterly by Fannie Mae and Freddie Mac. This version of the indices is based on a sample of repeat mortgage transactions for single-family residential properties in a national database of more than 17.5 million loans purchased or securitized by Freddie Mac and Fannie Mae between January 1975 and December 1992.[1] The house price indices are estimated by the repeat-sales method (Bailey, Muth, and Nourse 1963; Case and Shiller 1987, 1989; Shiller 1991), which can produce "constant-quality" house price indices even without detailed data on property characteristics because it uses information on the same physical units at two

* William Stephens is a Senior Economist at Freddie Mac. Ying Li is a Financial Economist at Fannie Mae. Vassilis Lekkas is a Senior Economist and Jesse Abraham is a Principal Economist at Freddie Mac. Charles Calhoun was Director of Economic and Demographic Studies at The Urban Institute and is now Senior Economist at the Office of Federal Housing Enterprise Oversight. Thomas Kimner was a Financial Analyst at Fannie Mae and is now a Financial Analyst at the Office of Federal Housing Enterprise Oversight. The research reported here was supported by Fannie Mae and Freddie Mac. The opinions expressed are those of the authors and do not necessarily reflect the positions of Fannie Mae, Freddie Mac, or their officers.

[1] The CMHPI was first officially released in the second quarter of 1994 by Fannie Mae and Freddie Mac. The official index was based on a sample of more than 4.5 million repeat transactions from the 22 million loans that Fannie Mae and Freddie Mac purchased or securitized from January 1975 to March 1994. The officially published index is a geometric repeat-sales index with a Goetzmann-like transformation.

points in time. The strengths and limitations of this method have been widely debated in the literature and are summarized below.

The next section reviews some of the major issues faced in valuation measurement based on price indices. This is followed by a discussion of the relationship between "ideal" indices and the geometric and arithmetic approaches to estimating repeat-sales indices. We then present a summary of the data and empirical results for the CMHPI. A number of cross-index comparisons are developed. This is followed by a discussion of potential sources of bias in the agency indices and a brief summary and conclusion.

Valuation Measurement and House Price Indices

All valuation indices involve the aggregation of individual property or attribute appreciation rates (see, for example, Diewert 1976). However, house price changes raise additional concerns. First, there is considerable heterogeneity among properties, so data must be aggregated over a large number of physical and locational attributes. Limits on information about specific markets may contribute to sampling or truncation bias. In addition, the variation in the information that is available can lead to measurement errors. Consequently, there can be a variety of price indices with very different derivations, all purporting to measure the same thing.[2]

Most consumer goods satisfy common, readily identifiable needs. Housing, in contrast, provides a heterogeneous collection of attributes, including privacy, shelter, comfort, proximity to employment, and access to public services. The price of housing depends on the demand and supply of these more fundamental attributes, which are embodied to a greater or lesser extent in old versus new construction, starter versus trade-up units, and even well versus poorly run government jurisdictions. As a result, price trends in one segment of the housing market may not be applicable to another.

Because of heterogeneity in property attributes, developing a house price index can require extensive data collection. Either all property types must be consistently represented in the sample used for analysis, or more narrowly defined, but consistent, indices can be calculated. Currently, property values are collected from a variety of sources, including property sales, appraisals, tax assessments, and owner estimates. Physical property attributes are sometimes available through broker listings, tax rolls, or government surveys. However, these data are not collected in a uniform way across time or by geographic location. Price index techniques therefore tend to be custom-designed to the data and are seldom broadly applicable.

Finally, the data that are available are often subject to considerable measurement error. Many believe that the highest quality measure of market value comes from an arm's-length purchase transaction. It may be difficult, however, to disentangle the value of other consumer durable goods transferred with the property, financing concessions, or time on the market, which can all be capitalized in the recorded price. Owner estimates

[2] Mark and Goldberg (1984) provide a good discussion of this issue. However, they do not recommend a particular index method, nor do they closely examine their repeat-sales results.

of property values tend to be biased upward, with a very wide variance compared with the arm's-length value (Goodman and Ittner 1992), while tax assessments are likely to be biased downward because of the possibility of owner challenges and infrequent updating (Clapp, Giaccotto, and Tirtiroglu 1990).

The selection of a particular method for calculating price indices inevitably reflects a balancing of collection costs against data and index quality. The hedonic model can be used in locations with "good" attribute data (Thibodeau 1989). Tax assessment data may also work in selected areas (Clapp 1990), while simply having a high volume of data may make it possible to use sample statistics such as the mean and median series in *Home Sales*, a publication of the National Association of Realtors (NAR).

We have adopted the repeat-sales method for the analysis of agency repeat transactions reported here. The mortgage transactions of Freddie Mac and Fannie Mae provide a readily available and efficient source of data for estimating repeat-sales indices. In addition to providing wide geographic coverage, the database of agency repeat transactions exists as machine-readable files and is easily updated in the normal course of business. The repeat-sales method is the only parametric approach that can be used to develop historical house price series using agency data. Until recently, Freddie Mac and Fannie Mae retained almost no information on property characteristics, precluding the possibility of adopting a hedonic approach based on their data.

In the following section, we consider the theoretical relationship between ideal indices based on geometric and arithmetic averaging, their relationship to repeat transactions, and the empirical approaches that have been proposed for estimating these ideal indices.

Ideal Indices and Repeat Transactions

A price index for period t is the ratio of average prices in period t to average prices in the base period, or period 0. If every transaction for which we have data were a repeat from period 0, there would be no need for regression analysis. In that case an ideal price index for a portfolio of houses could be constructed. Price indices for period t based on geometric and arithmetic means are given by

$$I_t^G = \left(\prod_{i \in J_t} \frac{P_{it}}{P_{i0}} \right)^{1/n_t} \tag{1}$$

and

$$I_t^A = \frac{\sum_{i \in J_t} P_{it}}{\sum_{i \in J_t} P_{i0}}, \tag{2}$$

respectively, where P_{it} is the observed price of unit i in period t, J_t is the set of units transacting in period t, and n_t is the number of units transacting in period t. These equations represent ideal indices in the sense that they embody perfect information about changes in the prices of the individual units in the portfolio since the base period. The repeat-sales method offers a way to construct an approximation to equations (1) and (2) in the absence of perfect information on base-period prices.

Geometric Repeat-Sales Estimators

Crone and Voith (1992) motivate the geometric repeat-sales approach from the definition of the growth rate of the value of a portfolio:

$$P_t = P_{t-s}\prod_{m=0}^{s-1}\left(1+\alpha_{t-m}\right), \tag{3}$$

where P_t is the average unit price in period t, and α_{t-m} is the average appreciation rate of the portfolio for period $t - m$, given by $(P_{t-m} - P_{t-m-1})/P_{t-m-1}$.

Bailey, Muth, and Nourse (BMN 1963) and Case and Shiller (1987, 1989) proposed estimators for $1 + \alpha_t$ based on the assumption that for individual house i, the price is a proportion c_i of the period average P_t, multiplied by a log-normally distributed disturbance u_{it}:

$$P_{it} = c_i P_t u_{it}. \tag{4}$$

Dividing equation (4) by the corresponding equation for period s, taking logarithms, and using data on repeat transactions for individual property units yields the empirical model first proposed by BMN:

$$Y_i = \sum_{\tau=1}^{T}\beta_\tau D_{i\tau} + \epsilon_i, \tag{5}$$

where $Y_i = \ln(P_{it}/P_{is})$; β_τ is a log price index ($\ln P_\tau$) for period τ ($\beta_0 = 0$); $D_{i\tau}$ is a dummy variable equal to 1 when $\tau = t$, –1 when $\tau = s$, and 0 otherwise; and random disturbance ϵ_i is distributed $N(0, \sigma^2)$. This model may be written in matrix notation for many repeat transactions as

$$\mathbf{Y} = \mathbf{D}\boldsymbol{\beta} + \boldsymbol{\epsilon}. \tag{6}$$

BMN assumed that the error vector $\boldsymbol{\epsilon}$ is independently and identically distributed with mean vector $\mathbf{0}$ and scalar covariance matrix $\sigma^2\mathbf{I}$, so that the ordinary least squares (OLS) estimator $\hat{\boldsymbol{\beta}}$ is the best linear unbiased estimator. The estimated house price index is given by

$$P_t = P_s e^{\hat{\beta}_t - \hat{\beta}_s} \tag{7}$$

for houses transacting in periods s and t and

$$P_t = P_0 e^{\hat{\beta}_t} \tag{8}$$

when $s = 0$. It is useful to rewrite the estimated OLS coefficient for period t as follows:

$$\hat{\beta}_t = \ln\left(\prod_{i\in J_t}\frac{P_{it}}{P_{is}e^{-\hat{\beta}_s}}\right)^{1/n_t}. \tag{9}$$

Computationally, the estimated coefficient for period t is the log of the ratio between the geometric mean of prices for units transacted in period t and an estimate of their geometric mean in the base period. This can be interpreted as an estimate of the geometric mean in equation (1) with the base-period prices estimated from initial transactions that occur in period s, for $s = 1, 2, \ldots, t-1$, because the base-period prices of most units are not observed.

Case and Shiller (1987, 1989) extend the BMN model by postulating that the error term in equation (6) is unlikely to be homoskedastic and depends on the length of the holding period between repeat transactions. In their model, the error term ϵ_i for repeat transaction i includes house-specific and random error components such that

$$\epsilon_i = H_{it} - H_{is} + N_{it} - N_{is}, \tag{10}$$

where H_{it} represents idiosyncratic variation in individual house prices, and N_{it} is a random disturbance term. Under the assumption that individual house prices are generated by a log-normal diffusion (Brownian motion) error process, we have

$$E\left(H_{it} - H_{is}\right) = 0 \tag{11}$$

and

$$E\left[\left(H_{it} - H_{is}\right)^2\right] = \left(t - s\right)\sigma_H^2. \tag{12}$$

H_{it} is assumed to be uncorrelated with ln P_t and with H_{jt}, $j = i$, for all t. N_{it} is distributed $N(0, \sigma^2{}_N)$ for $t = 0, 1, 2, \ldots, T$; uncorrelated with H_{it} and ln P_t for all i and t; and uncorrelated with N_{js} unless $j = i$ and $s = t$. This implies that ϵ_i is normally distributed with mean 0 and variance $(t - s)\sigma^2{}_H + 2\sigma^2{}_N$. Thus, this error structure assumes heteroskedasticity but no autocorrelation.

Case and Shiller (1987) propose a three-stage generalized least squares (GLS) estimation procedure. In the first stage, BMN's OLS procedure is applied. The second stage involves regressing the squared residuals from the first stage on a constant term and the interval between repeat transactions $t - s$ to determine the parameters of the Brownian motion and random error processes. The fitted values of the second-stage residuals are then used as weights in a third-stage GLS regression.[3] Following the terminology used by Shiller (1991), the resulting estimator will be referred to here as the Interval-Weighted Geometric Repeat-Sales (I-GRS) estimator.

The I-GRS estimator has the advantage of being based on a theoretical model for the relationship between individual house prices and the market average as described in equation (4). However, if what is ultimately desired is an estimator of the simple growth rate in the arithmetic mean value of prices of a portfolio of properties, the I-GRS index results in a downward-biased estimator because the geometric mean is always below the arithmetic mean when there is variation in the data.

Goetzmann (1992) proposes a simple adjustment to the I-GRS index that makes it more appropriate for portfolio valuation. The difference between the geometric and arithmetic means is an increasing function of the variation in the data. Goetzmann therefore suggests the following modified version of the I-GRS index:

$$P_t = P_0 e^{\hat{\beta}_t + \left(\hat{\sigma}_t^2/2\right)}, \tag{13}$$

where $\hat{\sigma}_t^2$ is an estimate of the variation of growth rates across properties at time t. This estimator is referred to here as the Interval-Weighted Linearized Geometric Repeat-Sales (I-LGRS) estimator. Like the I-GRS estimator, the I-LGRS estimator is closely tied

[3] To take heteroskedasticity into account in the second-stage regression, Case and Shiller divided the dependent variable by the holding period in that regression. Abraham and Schauman (1991) included a quadratic term in this regression to allow for the fact that the variance is unlikely to increase indefinitely with holding period. Their results support this hypothesis.

to an underlying model of house prices, but it more closely approximates the change in the arithmetic mean value of prices.

Arithmetic Repeat-Sales Estimators[4]

Shiller (1991) proposed several variations of arithmetic repeat-sales estimators. He defines them through the following regression model:

$$\mathbf{Z} = \mathbf{X}\boldsymbol{\theta} + \mathbf{u} \tag{14}$$

where $Z_i = P_{i0}$ ($i = 1, \ldots, n$) if the first sale occurred in period 0 and $Z_i = 0$ otherwise. The independent variable matrix **X** is obtained by replacing –1 in the dummy matrix **D** in equation (6) with the negative of the first transaction price and 1 with the second transaction price, leaving the rest of the matrix, all zeros, unchanged; θ is a vector of unknown regression parameters, and **u** is the error term.

Computationally, the OLS coefficient estimate of θ_t in equation (14) is the ratio of the least squares estimate of average price in the base period to the average price in period t, for houses transacted in period t, so that its reciprocal is an estimator of the arithmetic ideal index given by equation (2):[5]

$$\hat{\theta}_t^{-1} = \frac{\sum_{i \in J_t} P_{it}}{\sum_{i \in J_t} \hat{\theta}_s P_{is}}. \tag{15}$$

This estimator is referred to as the Value-Weighted Arithmetic Repeat-Sales (VWARS) estimator.

The error vector $\mathbf{u} = \mathbf{Z} - \mathbf{X}\boldsymbol{\theta}$ can be converted to a proportional error by dividing each row of **X** and **Z** by the price of the first sale corresponding to that row, thus weighting each house the same. Doing so produces the Equally Weighted Arithmetic Repeat-Sales
.9938 0S) estE609ncorrespo 0 0 6,mat Tf02167zer.mo ton u 1 Tf Tm(P)TjKtc R(11 Tw,)Tj-0

Conventional Mortgage Home Price Index

Data

An exchange of data between Freddie Mac and Fannie Mae resulted in a joint database of more than 17.5 million active or closed loans for the period from January 1, 1975, to December 31, 1992. Through a process of cleaning, sorting, and address matching, 1.57 million repeat transactions pairs were identified and used for further statistical analysis.[7]

The number of observations in the database is listed in table 1 for the nine U.S. census divisions and the whole United States, both the total number and those known to involve no refinancing transactions.[8] Nearly a third of the total sample and more than one-fifth of the purchase-to-purchase sample are from the Pacific division, primarily because of the importance of California mortgages in the portfolios of the two agencies. In addition, more than 80 percent of the observations in both samples come from five divisions, with the East and West South Central, Mountain, and West North Central contributing less than 20 percent.

Table 1. **Agency Repeat Transactions Data, 1975 Q1 to 1992 Q4**

	Sample Size	
Census Division	All Observations	No Refinancings
East North Central	293,246	48,649
East South Central	36,596	7,579
Middle Atlantic	205,559	35,990
Mountain	49,251	9,501
New England	145,382	18,739
Pacific	488,620	54,385
South Atlantic	215,015	41,940
West North Central	73,019	13,410
West South Central	63,389	14,367
United States	1,570,077	244,560

Median house prices in several data sets across four census regions are reported in table 2. The top panel reports homeowner estimates from the 1989 American Housing Survey, while the bottom panel reports homeowner estimates from the 1990 Census of Housing. These are compared with median price estimates from the NAR transaction sample, separate samples for all Fannie Mae and Freddie Mac transactions, and the combined agency repeat transactions sample.

[7] The loan-level database combines individual loan records from each agency. Properties eligible for inclusion were limited to single units with first liens based on conventional financing. In addition, a number of quality screens were imposed, and address fields were cleaned to remove extraneous characters. The data were sorted by ZIP code, street address, and transaction date, and consecutive records were then compared for an address match. Matching records were collapsed into a single property record containing the property values and transaction dates and written to a new data file.

[8] Refinancing transactions cannot be identified for most years prior to 1991 for Fannie Mae.

Table 2. **Median House Prices by Census Region ($)**

	Northeast	Midwest	South	West
1989				
American Housing Survey*	122,859	60,790	63,489	110,317
National Association of Realtors	145,200	71,300	84,500	134,900
Fannie Mae purchases	144,300	94,200	111,300	160,800
Freddie Mac purchases	135,000	88,000	112,500	149,900
Matched data set	155,600	101,500	125,000	163,800
1990				
Census of Housing*	124,400	62,300	65,800	126,200
National Association of Realtors	141,200	74,000	85,900	139,600
Fannie Mae purchases	145,000	100,000	117,000	167,100
Freddie Mac purchases	136,000	92,600	108,000	149,000
Matched data set	156,400	110,000	126,300	180,300

*Census and American Housing Survey figures are for owner-occupied housing units.

A comparison of the first two lines of each set shows that transactions tend to be of the higher valued properties in the existing housing stock. Agency purchases are a little higher still for a number of reasons: Many of the lowest valued properties will never trade but simply disappear from the housing stock; lower valued properties can more easily be purchased without mortgage financing and so may not appear in the agency databases; and the lower end truncation of the agency samples by government-insured (Federal Housing Administration and Department of Veterans Affairs) loans is more influential than the higher end truncation by the conforming loan limit.

Empirical Results

The agency repeat transactions data were used to estimate weighted repeat-sales indices by the I-GRS, I-LGRS, I-EWARS, and I-VWARS methods. Cumulative growth rates in house prices for the nine U.S. census divisions and for the whole United States were derived from the estimated indices (table 3).[9] In six of the divisions and for the U.S. average, the rank ordering of growth rates is the same: I-GRS, I-VWARS, I-LGRS, and I-EWARS, from lowest to highest.

In table 4 the same comparison is shown for two subperiods that had many more repeat transactions per quarter and exhibited quite different trends in house price growth. In the earlier period, which covers first quarter 1985 to fourth quarter 1990, some change in the ordering occurs, but the results are not as striking as for the later period, covering first quarter 1991 to fourth quarter 1992, for which growth rates were low as a whole. In the later period, I-VWARS fell below even I-GRS in every division except New England, suggesting that weighting by property values more than offsets the difference between the geometric and arithmetic means. One likely reason is that higher value units were affected more than lower value units by the slowdown in house price appreciation.

[9] Appendix table A.1 reports the historical index values that were used to derive the estimates of cumulative growth rates for the United States reported in table 3. Appendix table A.2 reports the historical index values for the nine census divisions based on the results for the I-LGRS index.

Table 3. **Regional and National Cumulative Percentage Change in House Prices, Agency Repeat Transactions, 1975 Q1 to 1992 Q4**

	Index Method			
Census Division	I-LGRS	I-GRS	I-EWARS	I-VWARS
East North Central	186.89	171.87	191.32	176.09
East South Central	139.18	122.91	139.34	125.49
Middle Atlantic	241.39	211.27	246.65	218.97
Mountain	186.17	166.72	197.80	188.99
New England	293.60	281.12	312.68	268.69
Pacific	441.14	426.01	471.45	437.06
South Atlantic	176.77	151.49	172.55	168.45
West North Central	115.96	98.93	124.58	99.75
West South Central	95.99	76.28	100.00	91.82
United States	216.63	196.78	224.59	205.55

Table 4. **Regional and National Cumulative Percentage Change in House Prices, Agency Repeat Transactions, Selected Periods**

	Index Method			
Census Division	I-LGRS	I-GRS	I-EWARS	I-VWARS
1985 Q1 to 1990 Q4				
East North Central	43.80	41.60	45.18	43.64
East South Central	24.52	21.65	24.81	22.13
Middle Atlantic	67.54	62.59	68.91	61.52
Mountain	11.07	8.84	12.85	9.83
New England	60.47	61.38	68.37	58.45
Pacific	81.53	81.41	86.36	83.58
South Atlantic	40.75	35.83	39.88	38.94
West North Central	22.31	18.89	21.81	19.27
West South Central	–5.27	–8.51	–4.47	–5.40
United States	46.24	43.82	48.10	44.84
1991 Q1 to 1992 Q4				
East North Central	7.75	7.31	7.86	6.78
East South Central	7.26	6.48	6.92	6.27
Middle Atlantic	4.23	3.29	4.14	2.50
Mountain	11.02	10.42	11.08	10.10
New England	–3.05	–2.41	–1.07	–2.24
Pacific	–0.45	–0.21	0.60	–1.01
South Atlantic	6.70	5.41	5.93	4.72
West North Central	6.42	5.44	5.85	5.27
West South Central	8.20	7.04	7.26	6.49
United States	4.22	3.68	4.39	3.18

The relationship between value-weighted and equally weighted indices does not appear to be a simple one. The balance of this article focuses on results for I-LGRS. This index has a clear relationship to the underlying model of house prices and is easier to compute than the arithmetic repeat-sales estimators. This focus is not meant to suggest that I-LGRS is the appropriate index for all uses. In particular, the arithmetic I-EWARS and I-VWARS indices are more precise for such applications as calculating the change in the mean value of a portfolio of housing units, while I-GRS would be more accurate for updating the values of individual housing units.

Cross-Index Comparisons

Notwithstanding statistical issues inherent in the estimation and application of repeat-sales estimators, the CMHPI is a reasonable approach and in some cases is superior to the existing alternatives. It is theoretically consistent with an underlying model of individual house price variation and is easily modified for estimating the change in the average value of a portfolio of properties. This is most obviously true for the valuation of properties securing loans purchased or securitized by Freddie Mac and Fannie Mae. However, comparisons of historical values of the agency index with the alternatives suggest that the agency index is a suitable indicator of trends in house price appreciation across the country.

Multiyear growth trends for the four U.S. census regions and the whole United States are shown in table 5 for the Census Constant-Quality Index, the NAR median existing price index, and the CMHPI (agency index).[10] At 10.1 percent, the agency index grew the slowest in 1977 to 1980, underperforming the other two indices in all four regions. Consumer price inflation, adjusted for mismeasurement of the housing component, rose only 9.2 percent annually during this period, less than any of the three indices.

Annual consumer price inflation during 1980 to 1985 was 5.5 percent, greater than all three house price measures. Alone among the three sources, the agency series shows it was the economically hard-hit Midwest that constrained national growth in house prices the most during this period, while all three measures indicate that the boom in the Northeast was already under way.

Even as general inflation slowed to 4.0 percent in the second half of the decade, the Northeast boom continued until 1987, while the West experienced a belated housing price boom in 1988–89. The Census Constant-Quality Index grew more slowly than consumer prices, largely because it missed the action in California prices. The greatest increase in house prices is recorded in the agency series.

Given the recent history of house prices across the country, increases of the magnitude reported in the NAR index for the United States as a whole appear implausibly high for 1992 and especially 1991, though the 12-month increase through October 1993 is a reasonable 2.5 percent. The Census Constant-Quality Index numbers look reasonable because they are low, but they have been low for years. The 1992 increase in the census

[10] The regional weighting used to create a national index is different for each of these three indices. The NAR index uses 1980 census weights, the census index uses the distribution of new construction in 1987, and the agency index uses 1990 census weights.

Table 5. **Alternative Measures of House Price Inflation, Annualized Growth Rates by Census Region (%)**

Index	Northeast	Midwest	South	West	United States
Census Constant Quality					
1977–80	11.3	10.3	12.6	16.0	13.1
1980–85	7.3	3.2	4.0	2.7	3.8
1985–90	6.1	4.0	2.4	4.9	3.9
1991	−3.8	2.7	1.9	0.8	0.8
1992	8.2	1.8	1.0	0.5	1.9
NAR median existing price					
1977–80	11.0	12.2	13.6	15.9	13.2
1980–85	7.9	2.6	5.2	1.4	4.0
1985–90	9.7	4.7	2.7	7.9	4.8
1991	0.5	5.1	3.5	5.4	5.0
1992	−1.3	5.0	3.6	−2.3	3.4
Agency, all observations					
1977–80	8.3	8.5	8.5	15.0	10.1
1980–85	10.4	2.1	4.6	3.6	4.9
1985–90	7.5	5.8	3.7	9.6	6.7
1991	1.6	4.8	5.0	2.3	3.4
1992	0.7	3.7	3.2	−0.2	1.7

index in the Northeast is difficult to reconcile with the behavior of the other indices. The agency numbers appear to conform to common knowledge in all regions and periods. In addition, the agency series measures the greatest deceleration in house prices relative to the 1980s.

Metropolitan-area price changes are reported in table 6, with graphs for selected cities in figures 1 and 2. The comparisons here are largely between the agency series and the familiar NAR series, although figure 2 also includes the appraisal-based series from the Real Estate Research Councils of Southern and Northern California.

Table 6. **Mean Quarterly Growth Rates (%) and Standard Deviations, Agency and National Association of Realtors Price Indices**

	Agency		NAR	
Metropolitan Area	Mean	Standard Deviation	Mean	Standard Deviation
Boston	2.0	2.8	1.8	4.6
Chicago	1.7	1.0	1.5	3.4
Dallas	0.5	2.1	0.5	3.2
Detroit	1.2	1.5	1.5	3.4
Houston	0.1	2.2	0.4	5.8
Los Angeles	1.7	2.1	1.5	3.7
Minneapolis	0.8	1.5	0.7	1.7
New York	2.2	2.5	2.0	4.1
San Diego	1.5	1.9	1.6	3.3
San Francisco	1.8	1.9	1.8	4.1
Washington	1.4	2.3	1.5	3.9

Figure 1. **Agency and NAR Index Comparisons**

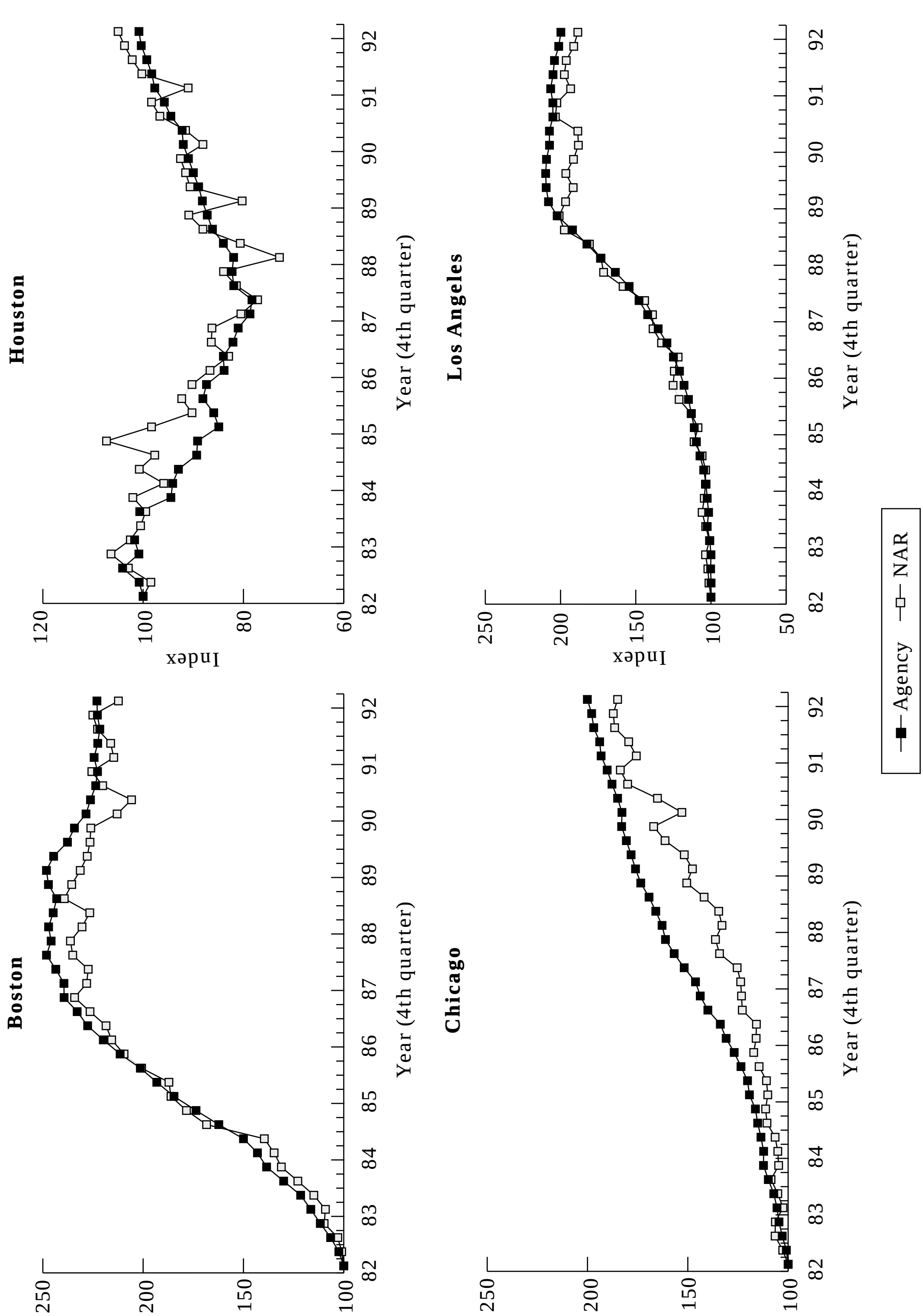
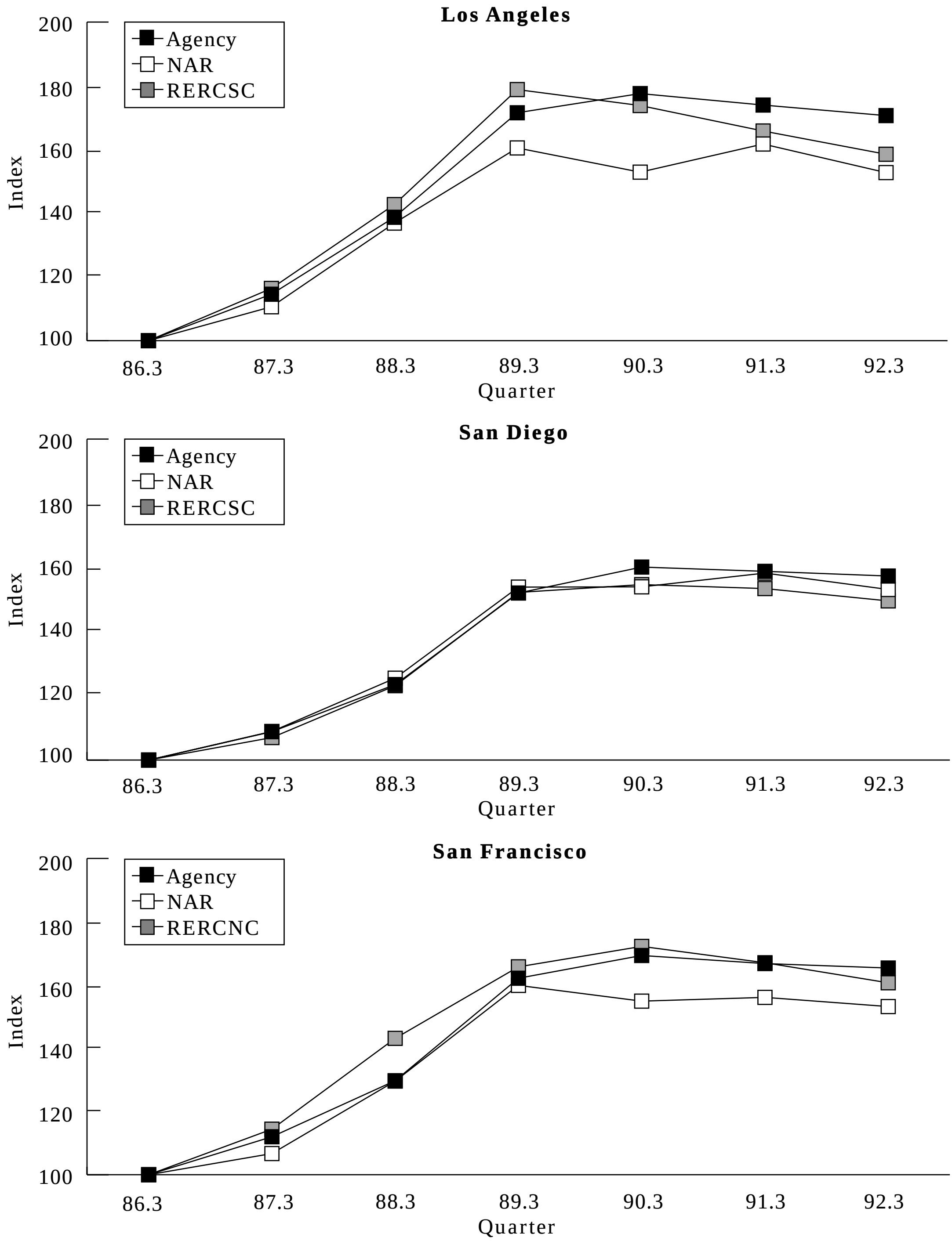


Figure 2. **Cross-Index Comparisons for Three California Metropolitan Statistical Areas**



Note: RERCNC = Real Estate Research Council of Northern California, RERCSC = Real Estate Research Council of Southern California.

It is evident from table 6 that long-term price appreciations for agency and NAR series measures are very similar. The largest differences are for Detroit and Houston, where the NAR mean growth rate is higher by 0.3 percentage point per quarter, or 1.2 percentage points per year. The most striking difference between these series is in their standard deviations: The NAR deviations are always larger—roughly twice those of the agency series.

A large standard deviation is not necessarily undesirable, since some variation is necessary to capture any movement beyond a mean growth rate. However, as shown in figure 1, the variability of the NAR indices for Boston and Houston appears to be more noise than information. The Chicago and Los Angeles panels show a somewhat different time-series pattern between the two measures, even though the difference in cumulative growth is modest.

For the three Pacific Division cities (figure 2), the strength of the agency indices relative to the NAR numbers likely reflects the mix of properties transacting: The NAR estimates may include the higher priced segment of the market, which softened more than the middle of the market, or perhaps fewer higher priced houses were included, resulting in a decline in the median price. The appraisal indices look at the same sample of properties every six months, which may not be representative of citywide trends. Despite these fundamental differences in the samples of properties, the three series produce reasonably consistent estimates of long-run house price inflation.

Recent studies by Case, Pollakowski, and Wachter (1991) and by Crone and Voith (1992) compared the properties of indices created using hedonic, repeat-sales, and transaction-sample means and medians.

Crone and Voith examined five methods for estimating house prices for identical samples: repeat sales, median and mean price indices, and constrained and unconstrained hedonic methods.[11] The accuracy of each method was judged both by how well it would have predicted the appreciation rates of individual out-of-sample houses and by its sensitivity to changes in sample size.

In the out-of-sample test, the three parametric methods (repeat sales, constrained hedonic, and unconstrained hedonic) dominated the mean- and median-based indices using a mean-squared prediction error (MSPE) criterion. Among the three parametric methods, the constrained hedonic method dominated repeat sales for a test of the null hypothesis of no difference in MSPE at the 10 percent significance level. Using the mean absolute prediction error, the repeat-sales method was more accurate than the others at the 5 percent significance level, with the two hedonic methods not significantly more accurate than the mean price index, and the median index was dominated by all other methods. The repeat-sales method was also the least sensitive to changes in sample size.

Case, Pollakowski, and Wachter (1991) compared repeat-sales, hedonic, and hybrid methods for constructing house price indices on the basis of a number of sample

[11] In the constrained hedonic method all trait prices increase at the same rate in a given year.

statistics.[12] They found that the hedonic model estimated on multiple transactions yielded lower average quarterly increases than hedonic estimates based on single transactions: 2.54 versus 2.81 percent over the 23 quarters estimated.[13] In addition, they found that the two repeat-sales models, with and without units that had undergone significant structural change, also resulted in lower average appreciation rates: 2.54 and 2.42 percent, respectively.

Comparing methods, the repeat-sales index estimated on unchanged units was found to perform best, with the exception of the hedonic model estimated on unchanged repeat transactions, in terms of both the width of the confidence intervals and the standard deviations of the disturbances. In addition, only the hedonic model estimated using repeat transactions on changed units had a higher squared correlation between the observed and predicted transaction prices than the repeat-sales method for unchanged units.

Potential Sources of Bias in CMHPI

Freddie Mac and Fannie Mae provide financing for a broad spectrum of low-, moderate-, and middle-income single-family properties. It does not immediately follow, however, that price indices generated from agency repeat transactions are applicable to the entire population of owner-occupied residential structures in the United States.

There are reasons that the properties in the agency repeat transactions sample may not be representative of all transactions or that transacting properties may not be representative of the entire stock of properties. The relevant question, however, is whether there are reasons to believe that any sample selectivity that does exist is correlated with price changes. The balance of this section considers four possible sources of bias in the agency index: renovation bias, transaction bias, refinancing bias, and revision volatility.

Renovation Bias

The CMHPI has been described as a “constant-quality” house price index because it uses information on the value of individual properties at two points in time. The index is truly constant quality only if properties retain the same physical attributes and those attributes are in fact valued the same by the market over time. Houses depreciate over time, both physically and relative to current fashion. Proper maintenance can prevent much of the physical deterioration, but the level of maintenance and number of major replacements required or performed is not a part of the regression. Similarly, additions and alterations may add value to the structure, but these are also not a part of the regression.

---

[12] Case, Pollakowski, and Wachter examined the average width of a 95 percent confidence interval around the predicted price, the estimated standard deviation of the disturbance term, the standard deviation of the quarterly increase in the index, and the squared correlation coefficient between actual and predicted transaction prices.

[13] They also estimated an average change of 1.99 percent for a subsample of changed repeats. However, this subsample was small enough that care should be exercised in interpreting that result.

For most purposes, an index measuring appreciation of actual existing houses, net of "expected" maintenance, is the relevant concept. The notion of a true constant-quality existing house is a theoretical construct that may be useful for defining and estimating structural equations of housing demand but is less relevant for applications such as marking a real estate portfolio to market or gauging changes in affordability. The repeat-sales method solves the problem with median and mean price indices of samples that differ from period to period, but it makes no artificial attempt to hold the quality of units constant.

The issue of renovations is less easily dismissed. The value of additions and alterations reported by the Census Bureau Series C-50 is roughly 0.5 percent of the value of all properties in the country.[14] Reflecting that figure, a constant 50 basis points could be subtracted from annual growth rates in all years (Abraham and Schauman 1991), or the adjustment could be permitted to vary over time with actual expenditures (Peek and Wilcox 1991). Of course, that approach assumes that all renovation costs add to value.

A level adjustment may capture the long-run value of renovations but could distort the timing of appreciation during the business cycle, particularly for small geographic areas that differ from the national pattern. However, review of the data shows that deflating the rehabilitation expenditures by local housing cost levels and the house price index removes most of the cycle in nominal rehabilitation spending (Abraham and Hendershott 1992). We conclude that timing differences across divisions are likely of second-order importance.

Transaction Bias

The source of data on repeat transactions is the ongoing business operations of Freddie Mac and Fannie Mae. As such, the specific loans and properties included in the sample are subject to historical agency purchase patterns and loan amount restrictions. For the future, with roughly 60 percent of all new originations purchased by one of the two agencies, there is a solid foundation for obtaining additional property matches.

The introduction of the Freddie Mac Weighted Repeat-Sales Index in 1990 raised the concern that the conforming loan limit, and the way it has increased over the years, would create serious measurement problems (Haurin and Hendershott 1991). These issues of bias were evaluated and discounted by Abraham and Schauman (1991, especially footnote 10) and found to be of little empirical significance (Abraham 1990). In an analysis of Fannie Mae repeat transactions, Calhoun (1991) found that the conforming loan limit had a negligible effect on the probability of sample selection.

Nonetheless, it is obvious that the conforming limit lowers the incidence of high-priced houses in the agency sample. Anecdotal reports from around the country suggest that price appreciation differs markedly by price segment: Most recently, the higher priced homes in the weak markets of California and the Northeast have seen much greater softness than mid-priced homes, as illustrated by table 4 and our earlier discussion of the value-weighted versus equally weighted indices.

---

[14] There are many recognized problems with the C-50 numbers; for example, they ignore the value of sweat equity.

The CMHPI measures appreciation in the middle of the market and specifically excludes appreciation or depreciation of the most expensive homes. This feature can be seen as a strength or a weakness, depending on how the index will be used.

Refinancing Bias

There are two conditions under which the inclusion of refinancing transactions could bias a derived price index: (1) if the successful refinancings are not a representative sample of all transactions or of the housing stock or (2) if appraisals fail to be an unbiased measure of value.

Appraisals are widely believed to contain significant measurement error; that is, they do not accurately predict the purchase price were the house to be sold. However, Dotzour (1988) does not find them to be statistically biased. Even if appraisers are unbiased, however, a consistent upward bias can still appear in closed refinancings if low appraisals are less likely than high appraisals to ultimately lead to loan closure and hence are not observed.

The correlation between appreciation and the ability to refinance is of special concern because many properties have fallen in value over the past several years. Owners who fail to refinance in the current low-mortgage-rate environment are more likely to be those facing an equity constraint under current underwriting standards. A price index heavily weighted toward refinancings could therefore have a greater incidence of rapidly appreciating houses than is present in the total stock.

Refinancing bias can be tested statistically. Consider an unrestricted form of the model described in equation (6):

$$\begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} D_1 & \beta_1 \\ D_2 & \beta_2 \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix}, \qquad (16)$$

where the subscripts 1 and 2 represent segmentation of the total data set into purchase-to-purchase repeat transactions and all other transactions, respectively. All periods are covered within each segment, so β_1 and β_2 have the same number of elements. It is then straightforward to test the null hypothesis that there is no significant difference between the indices from these two samples:

$$H_0 : \beta_1 = \beta_2. \qquad (17)$$

Using a standard F test, the unrestricted sum of squares comes from equation (13), where β_1 and β_2 need not be equal. The restricted sum of squares is from equation (6), which estimates a single coefficient vector with the entire data set of 1,570,077 observations. Data set 1 has 244,560 observations and includes only those matches where both transactions were purchases. Data set 2 is all remaining observations.

Test results for the joint data set over the entire 1975–92 sample period are presented in table 7. All the F ratios are significant at the 1 percent level, suggesting that the known purchase-to-purchase sample is different from the other matches in the joint data set.

Table 7. **Tests of Structural Change for Refinancing**

Census Division	F Value*
East North Central	4.19
East South Central	3.13
Middle Atlantic	5.02
Mountain	3.09
New England	2.00
Pacific	26.18
South Atlantic	9.51
West North Central	5.02
West South Central	4.96

*Values represent F statistic from test for structural change between entire data set and purchase-to-purchase data set. They are distributed $F_{71,\infty}$. All are significant at 1 percent level.

The results for the above tests are conservative, since Fannie Mae refinancings are not flagged before 1991. Therefore, what is actually being tested is the structural difference between the model in which both transactions are known to be purchases and that in which one or both transactions are refinancing or have an unknown purpose. On the other hand, this test is severe. The estimated coefficients represent the cumulative log changes in value since the base period of 1975. The F test challenges all of these coefficients to be statistically identical. Failure of the indices to satisfy this test does not necessarily disqualify all refinancing information.

The potential bias from using refinancing transactions must be weighed against the potential gain in index accuracy from regional disaggregation. Excluding refinancings cuts the size of the full sample by 84 percent and reduces the number of observations in two census divisions to fewer than 10,000 (table 1). The effects are magnified at lower levels of geographic disaggregation. For example, without using refinancing transactions, few MSA indices could be constructed with confidence.

Cumulative growth rates for the nine divisions and for the whole United States are listed in table 8 for two periods—from first quarter 1985 to fourth quarter 1990 and from first quarter 1991 to fourth quarter 1992—along with the difference in growth rates obtained when all known refinancings are left out of the sample. In the first period, dropping all known refinancings lowers the estimate of house price growth in seven divisions and raises it in two. In the second period, dropping all refinancings has a smaller negative effect, with only five divisions having lower growth rates when refinancings are excluded. In the Pacific during the later period, the rates are actually of different signs.

Nothing in the above analysis points conclusively to an optimal policy regarding the use of refinancing transactions. Excluding refinancings has noticeable effects, whose direction can vary over divisions and periods. Given the severe loss in sample size from restricting transactions to purchases, we decided to retain all repeat transactions.

Revision Volatility

As new loans are purchased or securitized, they can be used to generate additional repeat transactions. Since new loans can match with transactions from any earlier date in the

Table 8. **Impact of Refinancing Transactions on Cumulative Growth Rates, Agency I-LGRS Index, Selected Periods**

Census Division	All Observations	No Refinancings	Difference
1985 Q1 to 1990 Q4			
East North Central	43.80	44.79	0.99
East South Central	24.52	24.29	–0.23
Middle Atlantic	67.54	65.89	–1.65
Mountain	11.07	10.56	–0.51
New England	60.47	59.96	–0.51
Pacific	81.53	77.75	–3.78
South Atlantic	40.75	39.12	–1.63
West North Central	22.31	23.26	0.95
West South Central	–5.27	–9.99	–4.72
United States	46.24	44.87	–1.37
1991 Q1 to 1992 Q4			
East North Central	7.75	8.34	0.59
East South Central	7.26	7.75	0.49
Middle Atlantic	4.23	3.28	–0.95
Mountain	11.02	12.42	1.40
New England	–3.05	–3.29	–0.24
Pacific	–0.45	0.33	0.78
South Atlantic	6.70	5.68	–1.02
West North Central	6.42	5.82	–0.60
West South Central	8.20	6.15	–2.05
United States	4.22	4.07	–0.15

sample, the entire repeat transaction index is subject to revision as the new data are added and the index is reestimated.[15] The degree to which the earlier numbers must be revised can have a bearing on the value of the index for predictions and for calculating current loan-to-value distributions of mortgage portfolios. For census divisions in which few repeat transactions are observed, we would expect additional repeats to contribute useful information about house price appreciation, so we would have a relatively high tolerance for revisions in the index values. On the other hand, if the number of observed transactions is already large, we would expect the index values to be more stable.

We have examined how reestimating the index as new data are received changes the index values and the resulting rates of appreciation. For each of the nine census divisions, the I-GRS index was reestimated assuming that the sample ended in the fourth quarter of each year from 1988 to 1992. The resulting parameter estimates and covariances were then used to estimate annual percentage changes in house prices and asymptotic standard errors for each year from 1975 to the end of the series. The annual changes in the price indices were computed using the end-of-year (fourth-quarter) index values.

Results for the Pacific and East South Central Divisions appear in tables 9 and 10. Surprisingly, additional matches since 1988 have brought fairly large historical changes in the data-rich Pacific Division, especially in the final three periods of the series, which

[15] The agencies also purchase seasoned loans that may only create historical revisions without adding to the new period's information.

Table 9. **Sensitivity of Annual Percentage Change in I-GRS Home Price Index As New Data Are Added to the Sample of Repeat Transactions, Pacific Census Division, 1975–1992**

Period over Which Home Price Changes	Ending Date of Repeat Transactions Sample Used to Estimate I–GRS Home Price Index				
	88 Q4	89 Q4	90 Q4	91 Q4	92 Q4
75 Q4 to 76 Q4	19.34 (1.17)	19.22 (1.17)	19.41 (1.17)	19.48 (1.10)	19.52 (0.99)
76 Q4 to 77 Q4	26.12 (0.94)	26.50 (0.94)	26.73 (0.94)	27.04 (0.88)	27.27 (0.78)
77 Q4 to 78 Q4	16.79 (0.70)	16.30 (0.68)	16.24 (0.68)	15.92 (0.63)	16.14 (0.56)
78 Q4 to 79 Q4	17.25 (0.59)	17.62 (0.59)	17.27 (0.58)	17.53 (0.54)	17.61 (0.48)
79 Q4 to 80 Q4	13.60 (0.59)	13.38 (0.59)	13.51 (0.58)	13.08 (0.55)	12.67 (0.50)
80 Q4 to 81 Q4	8.13 (0.70)	8.09 (0.71)	8.17 (0.70)	8.25 (0.68)	7.77 (0.63)
81 Q4 to 82 Q4	–0.79 (0.60)	–0.72 (0.61)	–0.85 (0.61)	–0.95 (0.60)	–0.77 (0.56)
82 Q4 to 83 Q4	–0.27 (0.47)	–0.21 (0.48)	0.01 (0.48)	0.09 (0.47)	0.28 (0.43)
83 Q4 to 84 Q4	4.63 (0.41)	4.71 (0.42)	4.58 (0.42)	4.39 (0.40)	4.07 (0.37)
84 Q4 to 85 Q4	6.20 (0.28)	6.02 (0.28)	6.06 (0.28)	6.33 (0.28)	6.60 (0.26)
85 Q4 to 86 Q4	7.70 (0.20)	7.63 (0.20)	7.51 (0.20)	7.37 (0.19)	7.54 (0.17)
86 Q4 to 87 Q4	11.78 (0.28)	11.40 (0.25)	11.54 (0.24)	11.78 (0.22)	11.37 (0.19)
87 Q4 to 88 Q4	16.72 (0.35)	17.51 (0.31)	17.16 (0.29)	17.23 (0.25)	17.13 (0.21)
88 Q4 to 89 Q4		21.65 (0.28)	21.29 (0.26)	19.74 (0.21)	19.49 (0.17)
89 Q4 to 90 Q4			5.26 (0.24)	4.67 (0.19)	3.14 (0.13)
90 Q4 to 91 Q4				0.51 (0.16)	1.67 (0.11)
91 Q4 to 92 Q4					–1.19 (0.09)
Observations	102,301	131,343	163,605	258,352	488,620

Note: Asymptotic standard errors of predicted quarter-over-quarter changes in parentheses.

Table 10. **Sensitivity of Annual Percentage Change in I-GRS Home Price Index As New Data Are Added to the Sample of Repeat Transactions, East South Central Census Division, 1975–1992**

Period over Which Home Price Changes	Ending Date of Repeat Transactions Sample Used to Estimate I-GRS Home Price Index				
	88 Q4	89 Q4	90 Q4	91 Q4	92 Q4
75 Q4 to 76 Q4	2.06 (2.93)	2.08 (2.86)	2.33 (2.79)	2.47 (2.63)	2.46 (2.23)
76 Q4 to 77 Q4	9.17 (2.36)	9.28 (2.32)	9.43 (2.26)	9.72 (2.15)	9.99 (1.84)
77 Q4 to 78 Q4	9.41 (2.00)	9.85 (1.97)	10.09 (1.91)	10.49 (1.81)	11.03 (1.54)
78 Q4 to 79 Q4	4.59 (2.03)	4.38 (1.98)	4.54 (1.92)	4.57 (1.78)	4.51 (1.49)
79 Q4 to 80 Q4	3.40 (2.52)	3.74 (2.47)	3.95 (2.39)	3.17 (2.18)	3.15 (1.82)
80 Q4 to 81 Q4	2.42 (3.09)	2.68 (3.01)	2.53 (2.90)	3.26 (2.69)	3.50 (2.23)
81 Q4 to 82 Q4	–0.85 (2.63)	–1.01 (2.55)	–1.17 (2.46)	–1.00 (2.27)	–1.05 (1.88)
82 Q4 to 83 Q4	6.12 (2.17)	6.33 (2.12)	6.89 (2.06)	6.53 (1.90)	6.44 (1.58)
83 Q4 to 84 Q4	3.69 (1.87)	3.54 (1.81)	3.38 (1.75)	3.43 (1.61)	3.28 (1.33)
84 Q4 to 85 Q4	5.01 (1.14)	4.82 (1.10)	4.85 (1.06)	4.99 (0.98)	4.91 (0.80)
85 Q4 to 86 Q4	6.13 (0.79)	6.39 (0.76)	6.17 (0.73)	6.12 (0.67)	6.02 (0.55)
86 Q4 to 87 Q4	4.39 (1.17)	4.22 (1.04)	4.44 (0.96)	4.35 (0.82)	4.39 (0.61)
87 Q4 to 88 Q4	1.23 (1.43)	1.29 (1.25)	1.38 (1.11)	1.95 (0.90)	2.03 (0.64)
88 Q4 to 89 Q4		5.72 (1.29)	5.77 (1.12)	4.81 (0.86)	3.85 (0.54)
89 Q4 to 90 Q4			–0.53 (1.12)	–0.20 (0.81)	0.47 (0.45)
90 Q4 to 91 Q4				5.31 (0.68)	5.07 (0.39)
91 Q4 to 92 Q4					3.19 (0.32)
Observations	7,974	9,603	11,572	17,443	36,596

Note: Asymptotic standard errors of predicted quarter-over-quarter changes in parentheses.

have seen revisions up to several multiples of their standard errors. One reason for this instability is undoubtedly the severe drop in prices there, which prompted a lengthening of the time houses stayed on the market before the owners accepted prices below their original expectations. A second reason is the accuracy of the index itself in this data-rich division. The standard errors are quite low, so revisions represent new information with a high degree of certainty. The East South Central has historical revisions of the same general magnitude, but because the sample is smaller, fewer revisions exceed the standard errors of the initial estimates. Annual revisions tend to cancel each other when cumulative growth rates over several years are calculated.

The recent revision experience for the nine U.S. census divisions is summarized in table 11. The first column reports the cumulative revision of the fourth quarter 1987 to fourth quarter 1988 growth rate after four additional years of data (through fourth quarter 1992), the second column reports the change in the 1988–89 growth rate after three additional years of data (through 1992), and so on. The figures in parentheses are the standard errors of the initial estimates of house price growth.

With almost two-thirds of the revisions being negative, and with the negative revisions being of greater average magnitude than the positive changes, there seems to be a greater tendency toward downward than upward revision. There is also a clear cohort effect, with the 1988–89 change showing consistently large, negative revisions. However, without that year, the signs of the revisions are equally split, and more than half are below one standard error.

These revisions can be viewed as a positive feature of the index. The year 1990 was a transition period, when many areas around the country were undergoing sharp decelerations in house price appreciation. Initial sales likely did not fully reflect market weakness that was revealed by subsequent information.[16] A strength of the index is its ability to incorporate new information and to put it in its correct historical period, rather than forcing all previous errors into subsequent index values.

Summary and Conclusions

This article has introduced the CMHPI, based on 1.57 million repeats from a national database of 17.5 million loans purchased or securitized by Freddie Mac and Fannie Mae between 1975 and 1992. We have examined alternative repeat-sales methods and issues surrounding the representativeness of the agency repeat-sales database and compared the indices with figures published by NAR and the U.S. Bureau of the Census.

The repeat transactions method is appropriate for calculating price indices for such purposes as estimating changes in affordability or calculating the current value of a portfolio of housing units. Indices based on sample medians and means are subject to variation in the sample of houses being measured. Hedonic methods require information on individual property characteristics that is typically not available on a national scale, and they produce estimates that are not directly applicable to the valuation of a portfolio

[16] This revision experience is consistent with the notion that better performing properties trade more frequently than poorly performing ones. This “winners” theory is presented for the stock market by Shefrin and Statman (1985) and applied to house prices by Abraham and Schauman (1991).

Table 11. **House Price Growth Revisions and Standard Errors by Region and Years Subsequent to Preliminary Release**

Census Division	1987 Q4 to 1988 Q4	1988 Q4 to 1989 Q4	1989 Q4 to 1990 Q4	1990 Q4 to 1991 Q4
Years of revision subsequent to preliminary release	4	3	2	1
East North Central	–0.56	–1.38	–0.81	0.48
	(0.49)	(0.40)	(0.37)	(0.20)
East South Central	0.80	–1.87	1.00	–0.24
	(1.43)	(1.29)	(1.12)	(0.68)
Middle Atlantic	–0.63	–0.62	0.54	–0.06
	(0.59)	(0.45)	(0.43)	(0.36)
Mountain	0.09	–3.08	–0.71	–0.44
	(1.21)	(1.05)	(0.93)	(0.54)
New England	–0.05	–0.50	–1.54	0.38
	(0.62)	(0.48)	(0.49)	(0.38)
Pacific	0.41	–2.16	–2.12	1.16
	(0.35)	(0.28)	(0.24)	(0.16)
South Atlantic	–0.74	–1.15	–0.27	0.12
	(0.53)	(0.43)	(0.41)	(0.29)
West North Central	0.50	–1.61	0.28	0.27
	(1.21)	(1.02)	(0.88)	(0.48)
West South Central	1.85	–1.99	0.64	–0.02
	(1.07)	(1.03)	(0.96)	(0.69)

Note: Numbers in parentheses are standard errors of the initial estimates.

of existing properties. The repeat transactions method provides a solution midway between these two sets of problems.

Most of the statistical and data problems inherent in the repeat transactions method and the agency data are surmountable without excessive costs. The CMHPI is a cost-effective way to achieve both wide coverage and geographic disaggregation in the measurement of house price changes.

Appendix

Table A.1. **Agency Repeat Transactions Indices for the United States**

	Index Method			
Year and Quarter	I-LGRS	I-GRS	I-VWARS	I-EWARS
1975 Q1	100.00	100.00	100.00	100.00
Q2	102.56	102.44	103.96	103.29
Q3	102.68	102.45	103.06	102.52
Q4	105.80	105.43	106.53	105.96
1976 Q1	107.44	106.93	108.53	108.46
Q2	109.19	108.56	110.00	109.32
Q3	112.28	111.51	112.84	112.81
Q4	114.56	113.64	115.99	115.50
1977 Q1	116.51	115.45	117.03	116.92
Q2	120.39	119.16	120.85	120.79
Q3	124.00	122.60	124.12	124.23
Q4	128.61	127.02	128.97	129.13
1978 Q1	130.79	129.02	130.37	131.10
Q2	135.00	133.04	134.25	135.12
Q3	139.67	137.49	138.92	139.95
Q4	145.16	142.74	144.08	145.12
1979 Q1	148.00	145.38	146.70	147.11
Q2	152.08	149.23	150.73	151.04
Q3	157.32	154.20	156.08	156.74
Q4	160.60	157.25	159.88	159.92
1980 Q1	162.67	159.12	161.04	162.64
Q2	165.50	161.72	163.92	165.32
Q3	170.52	166.45	168.59	170.90
Q4	171.76	167.48	170.01	172.41
1981 Q1	173.96	169.45	171.97	174.78
Q2	176.95	172.19	174.73	177.63
Q3	178.75	173.76	176.66	179.50
Q4	179.50	174.31	176.76	180.08
1982 Q1	179.59	174.22	177.30	180.19
Q2	181.67	176.05	179.64	182.36
Q3	182.25	176.44	179.94	182.92
Q4	183.99	177.95	181.58	185.35
1983 Q1	186.51	180.20	184.17	188.05
Q2	188.47	181.91	185.82	190.11
Q3	190.00	183.20	187.32	191.64
Q4	191.68	184.63	188.86	193.26

Table A.1. **Agency Repeat Transactions Indices for the United States** *(continued)*

	Index Method			
Year and Quarter	I-LGRS	I-GRS	I-VWARS	I-EWARS
1984 Q1	194.78	187.45	192.00	196.46
Q2	198.14	190.48	195.54	200.20
Q3	200.17	192.25	197.70	202.12
Q4	202.77	194.57	200.23	205.04
1985 Q1	205.97	197.46	202.91	208.27
Q2	210.46	201.57	207.32	212.86
Q3	214.91	205.65	211.65	217.42
Q4	218.11	208.52	214.88	220.63
1986 Q1	223.02	213.01	219.51	225.39
Q2	227.88	217.45	224.02	230.38
Q3	232.80	221.96	228.81	235.62
Q4	237.70	226.44	233.88	240.87
1987 Q1	242.93	231.23	239.11	246.15
Q2	248.34	236.19	244.49	251.81
Q3	254.05	241.45	250.47	258.05
Q4	257.51	244.57	253.59	261.90
1988 Q1	262.71	249.33	258.91	267.37
Q2	269.32	255.41	265.59	274.25
Q3	273.45	259.18	269.33	278.75
Q4	277.11	262.49	273.02	282.87
1989 Q1	281.95	266.94	277.77	287.77
Q2	287.02	271.59	282.77	293.19
Q3	293.96	278.03	289.57	300.41
Q4	297.80	281.49	292.93	304.29
1990 Q1	299.69	283.10	294.11	306.40
Q2	300.63	283.80	294.36	307.62
Q3	302.03	284.95	295.43	309.12
Q4	301.21	283.99	293.89	308.44
1991 Q1	303.81	286.24	296.14	310.92
Q2	305.81	287.90	297.26	313.04
Q3	306.86	288.70	297.87	314.31
Q4	311.29	292.64	301.90	318.99
1992 Q1	312.62	293.67	302.71	320.34
Q2	313.54	294.32	303.24	321.54
Q3	315.06	295.53	304.52	322.90
Q4	316.63	296.78	305.55	324.60

Table A.2. **Agency Repeat Transactions Indices, I-LGRS Method**

	Census Division								
Year and Quarter	East North Central	East South Central	Middle Atlantic	Mountain	New England	Pacific	South Atlantic	West North Central	West South Central
1975 Q1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Q2	107.42	103.05	100.43	105.43	105.57	102.00	101.91	105.56	93.67
Q3	105.56	108.09	99.49	101.01	101.07	106.65	102.27	100.41	97.96
Q4	109.69	109.85	100.46	109.38	117.81	109.15	101.58	102.79	100.62
1976 Q1	113.18	113.07	102.87	105.25	103.18	111.79	102.81	108.56	105.10
Q2	114.51	109.23	101.77	113.99	108.75	119.05	105.74	108.26	101.62
Q3	115.14	110.50	107.38	117.72	113.95	127.44	105.84	106.67	106.30
Q4	120.83	112.98	105.78	118.92	120.08	131.03	104.57	111.41	108.75
1977 Q1	121.48	116.92	104.81	120.54	120.96	137.90	107.64	112.64	108.60
Q2	124.71	116.38	109.18	127.26	117.83	149.38	109.57	114.94	112.16
Q3	128.02	121.25	108.35	134.10	125.47	159.76	111.74	116.58	112.71
Q4	134.75	124.74	111.85	141.21	126.72	167.44	115.75	118.62	115.36
1978 Q1	135.57	128.16	112.90	142.42	135.16	172.38	115.29	122.30	117.23
Q2	142.92	124.55	114.79	150.92	133.18	179.54	118.65	125.72	123.26
Q3	147.74	133.84	120.12	158.54	140.17	187.53	120.18	129.02	123.08
Q4	155.34	139.04	120.63	163.38	147.60	195.18	125.19	131.66	131.32
1979 Q1	157.97	138.10	125.29	168.47	147.00	201.30	126.65	133.99	133.29
Q2	164.11	142.95	126.95	173.16	146.30	209.60	128.44	137.27	137.80
Q3	168.36	143.45	131.14	180.65	158.69	221.62	131.82	137.62	141.64
Q4	170.95	145.87	133.03	183.64	156.83	230.31	134.32	139.16	146.80
1980 Q1	171.51	150.80	132.87	187.25	151.04	237.68	139.12	142.73	143.61
Q2	172.06	145.04	132.95	192.65	151.62	245.78	142.20	143.59	153.03
Q3	173.78	150.99	143.38	194.99	161.78	255.31	144.19	145.91	155.21
Q4	173.42	151.05	141.08	200.69	163.17	260.25	147.77	148.59	152.48
1981 Q1	175.74	149.30	141.02	202.48	161.99	264.17	150.05	151.12	159.24
Q2	177.05	148.79	143.07	212.32	165.38	271.56	153.39	145.25	164.83
Q3	174.53	152.84	146.37	215.32	166.48	277.58	153.14	144.79	169.62
Q4	170.67	156.96	143.29	213.86	167.19	281.19	156.49	145.02	174.71

Table A.2. **Agency Repeat Transactions Indices, I-LGRS Method** *(continued)*

Year and Quarter	East North Central	East South Central	Middle Atlantic	Mountain	New England	Pacific	South Atlantic	West North Central	West South Central
	Census Division								
1982 Q1	170.64	153.38	144.33	213.07	172.86	275.84	157.15	148.77	176.83
Q2	170.46	154.27	149.81	221.46	173.60	278.09	158.74	145.31	180.90
Q3	172.56	151.13	148.62	220.98	173.94	278.48	159.95	146.97	182.64
Q4	173.90	155.92	149.76	227.55	176.45	279.65	161.20	147.83	183.44
1983 Q1	173.59	159.41	153.68	229.40	182.38	280.58	164.54	152.04	186.12
Q2	175.18	162.50	159.49	224.43	190.87	280.59	164.96	153.54	188.17
Q3	175.84	162.34	161.65	225.98	199.43	278.71	167.30	157.70	188.74
Q4	175.33	166.62	166.07	224.90	207.71	280.93	169.04	154.76	189.17
1984 Q1	177.03	167.38	170.50	228.13	217.75	286.10	172.75	157.10	187.32
Q2	178.58	169.44	179.62	223.61	227.67	288.22	174.82	160.25	190.85
Q3	178.88	170.22	183.60	224.48	236.15	292.00	176.64	161.21	189.80
Q4	181.20	172.77	190.67	227.59	242.14	292.80	178.65	162.88	188.82
1985 Q1	183.03	175.72	194.06	228.50	253.11	297.53	183.01	163.74	188.97
Q2	185.50	177.72	202.50	230.75	270.13	303.82	185.15	166.89	189.73
Q3	188.94	179.61	210.71	235.16	283.66	309.27	188.39	169.52	189.41
Q4	190.07	181.97	219.00	234.04	300.39	312.46	190.70	170.71	187.49
1986 Q1	193.46	185.12	224.09	239.36	313.24	317.62	196.04	172.42	192.44
Q2	196.95	187.97	233.57	241.95	327.11	322.66	198.75	174.39	196.19
Q3	200.84	189.49	246.54	242.57	344.55	329.23	201.52	177.04	192.37
Q4	204.59	193.70	257.48	242.43	359.31	336.26	204.50	179.26	190.95
1987 Q1	207.68	196.83	267.22	246.87	372.91	343.46	209.60	181.38	190.71
Q2	214.51	198.93	279.37	243.49	386.03	352.09	213.27	183.60	187.81
Q3	220.48	201.22	294.23	240.06	395.28	364.56	219.08	184.46	180.07
Q4	223.33	203.02	299.99	237.17	403.84	374.62	222.85	185.63	175.39
1988 Q1	227.29	205.49	306.88	240.15	413.54	388.39	226.95	187.02	174.02
Q2	233.07	207.67	316.01	239.53	421.99	402.39	233.97	188.15	177.53
Q3	236.81	208.65	319.22	241.45	421.61	421.27	237.89	188.95	172.42
Q4	238.41	207.98	321.44	238.67	425.16	438.77	240.93	190.48	171.63

Table A.2. **Agency Repeat Transactions Indices, I-LGRS Method** *(continued)*

	Census Division								
Year and Quarter	East North Central	East South Central	Middle Atlantic	Mountain	New England	Pacific	South Atlantic	West North Central	West South Central
1989 Q1	242.93	208.41	322.87	238.32	426.71	460.59	244.51	191.13	171.49
Q2	245.65	212.78	323.55	240.67	426.46	481.12	247.98	194.25	174.74
Q3	250.91	215.21	328.42	243.64	431.52	508.51	251.67	194.83	176.03
Q4	253.03	216.88	330.60	244.28	432.58	524.06	254.64	196.52	176.93
1990 Q1	256.01	217.59	330.41	245.57	426.36	532.80	255.88	198.71	177.04
Q2	259.78	218.12	328.89	249.58	418.00	535.99	256.84	198.66	178.10
Q3	262.30	218.90	327.88	252.73	414.80	540.52	258.15	199.95	179.04
Q4	263.20	218.80	325.14	253.80	406.18	540.10	257.59	200.27	179.01
1991 Q1	266.25	223.00	327.52	257.75	406.00	543.60	259.38	202.93	181.14
Q2	270.42	225.53	329.21	262.01	400.09	543.10	262.94	204.66	183.72
Q3	273.34	226.53	329.04	263.77	397.86	544.10	263.52	206.08	185.41
Q4	276.26	230.84	335.49	270.11	399.64	548.50	269.45	209.35	188.60
1992 Q1	278.25	232.71	337.97	274.28	397.00	547.37	271.19	210.30	190.46
Q2	282.27	234.75	335.77	277.92	393.27	545.54	272.84	212.72	192.30
Q3	283.73	237.51	339.32	280.57	393.87	543.18	275.05	213.82	194.70
Q4	286.89	239.18	341.39	286.17	393.60	541.14	276.77	215.96	195.99

References

Abraham, Jesse M. 1990. Statistical Biases in Transactions-Based Indexes. Mimeo. Freddie Mac.

Abraham, Jesse M., and Patric Hendershott. 1992. Patterns and Determinants of Metropolitan House Prices, 1977–91. Working Paper No. 4196. National Bureau of Economic Research.

Abraham, Jesse M., and William S. Schauman. 1991. New Evidence on Home Prices from Freddie Mac Repeat Sales. *AREUEA Journal* 19(3):333–52.

Bailey, Martin J., Richard F. Muth, and Hugh O. Nourse. 1963. A Regression Method for Real Estate Price Index Construction. *Journal of American Statistical Association* 58:933–42.

Calhoun, Charles A. 1991. Estimating Changes in Housing Values from Repeat Transactions. Paper read at the Western Economic Association International meetings, June 30–July 3, Seattle.

Case, Bradford, Henry O. Pollakowski, and Susan M. Wachter. 1991. On Choosing among House Price Index Methodologies. *AREUEA Journal* 19(3):286–307.

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

Case, Karl E., and Robert J. Shiller. 1989. The Efficiency of the Market for Single-Family Homes. *American Economic Review* 79(1):125–37.

Clapp, John M. 1990. A Methodology for Constructing Vacant Land Price Indexes. *AREUEA Journal* 18(3):274–93.

Clapp, John M., Carmelo Giaccotto, and Dogan Tirtiroglu. 1990. Repeat Sales Methodology for Price Trend Estimation: An Evaluation of Sample Selectivity. Working Paper 90-16. University of Connecticut School of Business Administration.

Crone, Theodore M., and Richard P. Voith. 1992. Estimating House Price Appreciation: A Comparison of Methods. *Journal of Housing Economics* 2(4):324–38.

Diewert, W. Erwin. 1976. Exact and Superlative Index Numbers. *Journal of Econometrics* 4: 115–45.

Dotzour, Mark. 1988. Quantifying Estimation Bias in Residential Appraisal. *Journal of Real Estate Research* 3:1–11.

Goetzmann, William. 1992. The Accuracy of Real Estate Indexes: Repeat Sales Estimators. *Journal of Real Estate Finance and Economics* 5(1):5–53.

Goodman, John L., Jr., and John B. Ittner. 1992. The Accuracy of Home Owners' Estimates of House Value. *Journal of Housing Economics* 2(4):339–57.

Haurin, Donald R., and Patric H. Hendershott. 1991. House Price Indexes: Issues and Results. *AREUEA Journal* 19(3):259–69.

Mark, Jonathan H., and Michael A. Goldberg. 1984. Alternative Housing Price Indexes: An Evaluation. *AREUEA Journal* 12(1):30–49.

National Association of Realtors. *Home Sales*, various issues.

Peek, Joe, and James A. Wilcox. 1991. The Measurement and Determinants of Single Family House Prices. *AREUEA Journal* 19(3):353–82.

Shefrin, Hersch, and Meir Statman. 1985. The Disposition to Sell Winners Too Early and Ride Losses Too Long: Theory and Evidence. *Journal of Finance* 40(3):777–90.

Shiller, Robert J. 1991. Arithmetic Repeat Sales Price Estimators. *Journal of Housing Economics* 1(1):110–26.

Thibodeau, Thomas G. 1989. Housing Price Indexes from the 1974–1983 SIMSA Annual Housing Surveys. *AREUEA Journal* 17(1):100–17.

U.S. Bureau of the Census. Various years. *Current Construction Reports: Expenditures for Residential Improvements and Repairs*. Series C-50.