FHFA WORKING PAPERS

Working Paper 10-1

Estimating Median House Prices

Andrew Leventis, Senior Economist

Office of Policy Analysis and Research
Federal Housing Finance Agency
1700 G Street NW
Washington, DC 20552
(202) 343-1502
Andrew.Leventis@fhfa.gov

February 2010

The author gratefully acknowledges the comments of Patrick Lawler, Robert S. Seiler, Jr., and Robert Collender at FHFA.

Estimating Median House Prices

Abstract

Nondisclosure laws and other factors have hindered the production and release of median home price in many areas across the country. This paper attempts to fill the gaps and develops a simple approach to estimating median prices for a geographically complete set of areas. The methodology begins by aggregating mortgage-level data from Fannie Mae, Freddie Mac, the Federal Housing Administration, and First American CoreLogic. Redundant observations are removed from the aggregate data and median home prices are calculated from the resulting dataset. Using county recorder data obtained for several hundred counties across the United States, the methodology then estimates the relationship between the calculated median prices and medians determined from the recorder data. The paper illustrates how this relationship can then be used to extrapolate median home values for areas for which recorder data are unavailable. Because one potential application of this approach would be to support the setting of conforming loan limits in high-cost areas, which are calculated as a function of median prices, actual 2009 high-cost-area conforming loan limits are then compared to the loan limits that would have been in effect had the approach been used to set the limits. The paper also briefly discusses the application of the basic methodology to estimating a national average home price, a statistic important for the setting of the national conforming loan limit.

BACKGROUND

Although house price medians are regularly produced for some of the largest metropolitan areas in the United States, no single source provides a geographically complete set of medians for the entire country on a regular basis.[1] The Federal Housing Finance Agency (FHFA) produces house price indexes (HPI) that reflect price *changes* for virtually every county in the United States but price *level* information (such as average prices or median prices) currently is not released. The data sample used by FHFA to construct its indexes is constrained to homes financed with conforming mortgages. This constraint could introduce bias into any median or average price measure that might be produced.[2]

To be sure, broader theoretical problems plague any summary measure of home values that might be constructed by FHFA or any other entity. For example, sample selection bias can be a significant problem if measurement of values only occurs when a specific event happens. When transaction prices (e.g. sales prices) are used to measure median home values, for instance, that metric can be biased if homes that transact in any given period are not representative of the underlying housing stock.

Theoretical issues notwithstanding, the absence of a reliable, geographically complete series of value metrics frustrates academics and industry observers. Median prices measures are needed, for example, to study cross-sectional differences in housing affordability. Also, the Federal Housing Administration (FHA) and FHFA require local median prices (values) to set local conforming loan limits, which Congress has mandated to be set as a function of median prices.

This paper will not address the problem of the theoretical divergence between median transaction prices and the median home value for the overall housing stock. Rather, the paper will be more pragmatically oriented; it will attempt to produce a reliable methodology for estimating median transaction prices for homes across a geographically complete set of areas in the United States. The target of estimation will be the median *sales price* and the focus of attention will be on developing a single methodology that can employ available mortgage data to estimate medians for every metropolitan area in the country.

Ad Hoc Approaches to Estimating Median Prices

In the last three rounds of loan limit determination,[3] FHFA and FHA have relied on an improvised methodology developed by HUD to estimate local median prices for all 3,200

---

[1] The National Association of Realtors (NAR) publishes the most widely-cited regular series. For the first quarter of 2009, 159 metropolitan areas were covered in its quarterly publication of median prices. Median prices can also be calculated from American Community Survey. Those data are available for a larger number of areas, but are still geographically incomplete and become available with a very significant delay.

[2] The direction of the bias is unclear. Because very expensive homes often require jumbo-sized mortgage financing, the data sample does underreport sales transactions for expensive homes. At the same time, homes not requiring any financing (cash sales) and homes financed with government-backed mortgages, transactions that presumably would tend to involve less expensive homes, are also not included.

[3] Median prices were determined for the setting of temporary jumbo-conforming loan limits under the Economic Stimulus Act of 2008. Later in 2008, 2009 loan limits were set using new estimates of median prices. The 2009

counties and county-equivalents in the United States. The improvised approach was necessary because transactions data are not available from all county recorder offices in the United States. In some cases, the absence of data stems from local non-disclosure laws that forbid the release of transaction prices. In other cases, local data are not available because none of the private data vendors that collect county-level sales data has established data collection programs with the specific counties. These counties tend to be small, where vendors find that the benefits associated with collecting the data (i.e., the additional revenue from resale of data) is small relative to the fixed costs associated with implementing a data collection program.

The HUD methodology used sales transaction data from county recorder offices, where they were available.[4] Where they were not available, data from the National Association of Realtors, which are based on MLS listings information, were used. In areas where those prices were not available, the last-resort option was generally to use home values from the American Community Survey (ACS) or the 2000 Census. Self-reports of home values from homeowners formed the basis for median price calculations in both cases. Values were taken from respondents who recently moved where statistical precision was sufficient; otherwise, median values were calculated from all respondents' indications of home value. In all cases, the estimates of value were stale and needed to be "brought forward" to a more recent period. FHFA's metropolitan area or state non-metropolitan price indexes were used for converting old median values for specific areas to more recent values.[5]

The *ad hoc* ACS-based approach was not the first methodology explored by HUD.[6] In the early part of this decade, HUD financed research into an approach that leveraged Home Mortgage Disclosure Act (HMDA) data. The research, conducted by Bob Cotterman of Unicon Research and Charlie Calhoun (Urban Institute, Calhoun Consulting), began by using data from Fannie Mae and Freddie Mac, FHA, and VA mortgage-level data to estimate the relationship between borrower and loan characteristics and house prices. Once those relationships were established, house prices were estimated for every mortgage represented in the HMDA data. For each agency's loans in HMDA, distributions for the model-estimated prices were then compared against distributions for the actual prices to assess accuracy. Ultimately, after testing six variants of the basic methodology, the research found no dominant methodology that produced superior results in fitting existing data.

Using the six different methodology variants applied to the mortgage-level HMDA data, the analysis then estimated year-specific median prices for different metropolitan areas and compared those to externally-sourced median home values. The external medians used for comparison purposes included those produced by Case-Shiller-Weiss, the National Association of Realtors (NAR), and others constructed from the American Housing Survey, the Mortgage Interest Rate Survey, and the 2000 Census. Data restrictions prevented the

---

limits were set under the terms of the Housing and Economic Recovery Act of 2008. 2010 loan limits were recently set under the terms of PL111-88.

[4] These median prices were supplied by RadarLogic, which calculated medians using sales transaction data it sub-licensed from one of the few major suppliers.

[5] Those data can be downloaded at: http://www.fhfa.gov/DataTools/Downloads/Pages/House-Price-Index.aspx.

[6] FHA in fact has been in need of localized price medians for many years because its loan limits have been tied to area price levels.

authors from testing the statistical significance of the difference between the HMDA-estimated values and the externally-sourced medians. Accordingly, the analysis focused on determining whether the estimated medians fell close to or within the range of the externally-sourced median home values.

Because HMDA data are released with a very significant delay, the contemplated HMDA-based approach was plagued by an unavoidable lack of timeliness. HMDA data were and continue to be released with at least a nine-month lag; transaction-level data from 2008, for example, were released in September 2009.[7] The HMDA-oriented methodology thus was and continues to be unattractive for any application that requires the estimation of recent median prices.

NEW METHODOLOGY

Basics

This paper will use neither the HMDA-based approach nor the piecemeal methodology employed by HUD in recent loan limit determination. Rather, the basic approach involves estimating the relationship between medians derived from available county-recorder data, which are assumed to reflect "true" medians, and medians calculated using home values from mortgage-level data. Importantly, the latter are available for homes throughout the United States, with no significant coverage deficiencies. The empirical relationship between the two medians is then estimated using data from counties where there is "overlap"—where both datasets have coverage. Using the estimated relationship (the coefficients of a regression model) in the overlap areas, the approach then can estimate median homes prices in areas where county-recorder data are unavailable.

To assemble a mortgage-level dataset containing home prices for a geographically complete set of areas, FHFA has combined mortgage data from Fannie Mae and Freddie Mac (the "Enterprises"), FHA, and First American CoreLogic. The Enterprise data arrive at FHFA on a quarterly basis and include house prices for homes financed with mortgage bought or securitized by Fannie Mae or Freddie Mac since the 1970s. The FHA mortgage data, which are supplied to FHFA each quarter through a data-sharing arrangement with HUD, include house prices for homes financed with FHA-insured mortgages since 1975.[8] The First American CoreLogic information, known as the "LoanPerformance" (LP) dataset, reflects historical sales prices for homes financed with securitized subprime and jumbo mortgages.[9]

County-recorder-based median prices, which are referred to as "unrestricted" medians because homes sold with every type of financing (including cash sales) are included, are available for several hundred U.S. counties. FHFA licenses transaction data from several

[7] See http://www.ffiec.gov/hmcrpr/hm091108.htm.
[8] Reverse mortgages are excluded.
[9] Other commonly used sources of mortgage-level data from loan servicers could also be used to augment the pooled dataset. For example, assuming permission was granted for use of its data, information from Lender Processing Services (LPS)—sometimes known as the "McDash" data—could theoretically augment the pooled series.

hundred counties from DataQuick Information Systems ("DataQuick"). Medians based on the pooled (Enterprise+FHA+LP) dataset are referred to as "financing-restricted" medians.

Duplicate Removal

The pooled dataset comprised of the three mortgage-level data sources is extraordinarily rich. Home sales included in the data should have full geographic coverage across essentially every county in the United States[10] and will include sales prices for homes financed with many of the most common types of financing. The gaps in the dataset include homes purchased with cash, VA-financed mortgages, unsecuritized subprime and jumbo mortgages, and other conventional mortgages held in portfolio. In the current market environment, where Enterprise- and FHA-guaranteed mortgages are the dominant forms of financing, the financing-related coverage gaps will not be particularly large.

One challenge associated with working with this pooled dataset involves removing duplicates. The Enterprise and LP datasets have no duplicates because the Enterprises were not allowed to purchase or securitize jumbo mortgages and the securitized subprime loans in LP were, as a matter of course, not guaranteed by the Enterprises. A handful of FHA mortgages appear in the LP and Enterprise datasets, however.[11]

Removing duplicates is an imperfect exercise, fraught with the potential for removing too many duplicates or removing too few. Fortunately, the fundamental imprecision of the exercise is not particularly pernicious in this context. To be sure, it would be reasonable to assume that median prices in the restricted data sample would be better predictors of full-market medians if the duplication removal algorithm were perfect. At the margin, however, error in the duplication removal process (i.e., excessive or insufficient removal) should have a modest impact on overall predictive accuracy. Also, the impact should be small by virtue of the relative paucity of FHA loans in the Enterprise and LP datasets.

To identify "duplicates" in the multiple data sources, mortgages were identified that: were collateralized by homes in the same zip code, had loan amounts within one percent of one another, had reported interest rates within one-quarter of a percentage point, and closed within 31 days of each other. This rather parsimonious set of matching variables is likely prone to excessive removal of duplicates.[12] Given that the underlying data are stored in different data systems, each with its own idiosyncratic data storage nuances, however, it seemed reasonable to employ this liberal tolerances-based approach.

---

[10] The Enterprise dataset alone contain 2008 transaction data for more than 3,000 counties and county-equivalents in the 50 states and Washington, D.C. This reflects about 96 percent of the 3,141 counties in the country (excluding territories).

[11] If additional loan servicer data, such as the McDash dataset, were added to pooled datasets, the extent of overlap would be much greater.

[12] Undoubtedly, cases will exist where: (a) different transactions have similar loan attributes and (b) the distinct transactions from (a) are found in different datasets. In these cases, the redundant transactions will be assumed to be "duplicates" and thus removed when they are in fact separate transactions.

Data Restrictions and "Area" Definition

Home prices for single-family dwellings are the target of estimation in this exercise; prices for condos, coops, and other property types are removed from all datasets prior to model estimation. Because the share of transactions accounted for by each property type can vary significantly from period to period (particularly for small geographic areas), pooling transactions from different property types can render "median" measures of dubious value. With pooling, period-to-period changes in median values reflect a mix of changes in home values as well as changes in the mix of properties that transact in the respective periods.

The estimation approach proffered in this analysis focuses exclusively on data from first-lien mortgages. Second-lien mortgages introduce potential redundancy into the mortgage-based transaction dataset; i.e., the same transaction could be reflected in multiple mortgages. Given that that redundancy might bias into median price estimates,[13] second-lien mortgages were removed before the financing-restricted medians were calculated.

Because one of the primary potential uses for median prices involves loan limit determination, this paper estimates median prices for the geographic aggregations that are currently used for loan limits. For the purposes of loan limits setting, "areas" include metropolitan statistical areas (MSAs), micropolitan statistical areas (µSAs), and counties that are not in either MSAs or µSAs ("unaffiliated" counties). There are 366 MSAs and 574 µSAs in the United States, excluding those in Puerto Rico. With 1,355 unaffiliated counties, the total number of "areas" requiring median estimates is thus 2,295.

It should be noted that the "median price" used to determine current loan limits is currently set equal to the median price in the highest-cost component counties for metropolitan and micropolitan areas comprised of more than one county. Statute requires that it calculate medians in this fashion. FHFA, although under no such mandate, opted to do so to align its median prices (and thus loan limits) with those of FHA.

The high-cost-county approach is not theoretically appealing because it significantly distorts median price estimates in some areas. It also requires the calculation of county-specific median prices in metropolitan and micropolitan areas. In this analysis, owing to the theoretical problems and the peculiarity of the "median" definition under the high-cost-county rule, estimating medians under the standard definition is the primary focus. As one might expect, and as is confirmed empirically, the relative attractiveness of the estimation approach delineated here does not differ significantly if the high-cost-county paradigm is applied.

For a number of metropolitan and micropolitan areas, the county recorder data available to FHFA were geographically incomplete. For example, recorder data at times were not available for all counties within a given metropolitan area. The incomplete geographic coverage significantly hinders the ability of the financing-restricted median price (which has full geographic coverage) to explain the unrestricted median in certain areas. Accordingly,

---

[13] Home values for houses financed with simultaneous first- and second-lien mortgages might systematically differ from values for other homes.

when isolating "overlap" areas for which a regression model should be estimated, these areas are omitted.

Other exclusion rules are also applied in an attempt to improve the estimation. In some cases, for example, a small number of sales transactions were available in counties for which the licensed dataset was not supposed to have coverage. No median prices are included for those areas. Also, the estimation model was not fed any median prices where the underlying sample size is ten or fewer observations.

RESULTS

Coefficient Estimates

The initial specification for the estimation model is simply:

$$\text{Median Price Full Market}_{a,t} = \beta_0 + \beta_1 (\text{Financing Restricted Median Price})_{\text{a,t}} + \varepsilon_{a,t}$$

Each observation is a median price for a given area, *a*, for homes financed with mortgages closed in a given year, *t*. The underlying data are a panel dataset, where year-specific median prices are available for each area. Median prices are calculated for the years 2002 through 2008, although some areas do not have the entire 2002-2008 series.

β_1 reflects the extent to which the estimated full-market median is estimated to increase for every dollar increase in the restricted median. *A priori*, one might expect that the coefficient to be near one, although theoretical arguments could be made for greater or lesser values.

Column (1) in Table 1 shows summary statistics and coefficient estimates from this simple regression.[14] With a statistically significant β_1 value of 1.11 and an R-Squared value of .891, the regression clearly shows that the financing-restricted median price closely tracks the full-market medians. The sample size of 1,552 observations, given up to seven years of data for each area, reveals that an average of about 222 areas per year are included in the basic data.

Column (2) explores a simple time-series nonlinearity in the relationship between the two median price measures; the specification allows the relationship between the medians to differ for years 2007 and after. Given the structural market changes that occurred over the latest years, it is reasonable to give the model this added flexibility. The statistically significant coefficient reveals that that flexibility produces added explanatory power, with the marginal effect of increases in the restricted median prices being significantly larger in 2007 and 2008.

The significantly larger coefficient in recent years would seem to be a function of the growing contribution of FHA loans to the pooled restricted dataset. The share of FHA transactions in the pooled data grew substantially over the last two years. Given that FHA-financed homes

[14] The regression model is estimated by ordinary least squares without heterogeneous observation weights. Theoretical arguments could be made both for and against weighting each observation by the number of transactions used in the underlying median calculations. The qualitative results presented in this paper do not differ substantially if weighted least squares is used, however.

tend to be toward the lower end of the price distribution, this growing share would tend to lead to a growing divergence between the restricted median and the full-market median.

Column (3) provides for differences in median relationship across the price spectrum. Dummy interactive variables are added for restricted median values about $200,000 and $400,000. The variables, *Median_price_above200* and *Median_price_above400*, take values of zero for restricted medians below $200,000 and $400,000 respectively. For values above the respective threshold levels, the variables are set to the restricted median prices.

The interactive variables provide the model with flexibility in the event that, at higher median prices, additional increases in the restricted median are associated with smaller or larger increases in the unrestricted medians. This situation might arise, for example, because the mortgage-level (restricted) dataset has imperfect coverage for homes financed with jumbo mortgages.[15] Because of the limited coverage, it may be very difficult to obtain extremely large median values in the restricted data. For higher values of the pooled median, each incremental increase in that median might only occur with significant growth in the full-market medians.

With the interactive variables, the estimate for the total incremental impact of increases in the restricted median is the sum of the respective coefficients. For median restricted prices below $200,000, the estimated marginal effect is merely the coefficient on the *Median_price_restricted* variable. For restricted medians between $200,000 and $400,000, the incremental impact is that coefficient plus the estimated coefficient on the *Median_Price_above200* variable. For restricted median prices above $400,000, the aggregate estimated impact is that sum plus the *Median_Price_above400* coefficient.

As shown in Column (3) of Table 1, the regression finds strong evidence of cross-sectional nonlinearities in the relationship between the unrestricted and restricted medians. The coefficients on the interactive variables are statistically significant at conventional levels. The signs of the coefficients are both negative, however, indicating that the empirical evidence is not consistent with the example described above (higher restricted medians are associated with particularly large unrestricted values because of the loan limits). While for low-cost areas, each dollar increase in the restricted median raises the predicted unrestricted median by $1.23, the marginal impact of increases in the restricted median is much lower (about $1.12) for the most expensive areas.

Specification (4) adds geography-specific information by including a regressor that reflects the median house price in the closest nearby area. For every area (MSA, µSA, or unaffiliated county), distances to every other area are calculated and the geographically closest neighbors are found.[16] After determining whether unrestricted median prices are available for those closest neighbors, the methodology then defines the "nearby median" as the median price for the closest neighbor that has an unrestricted median. The use of the unrestricted medians

[15] The limited coverage, as indicated earlier, stems from the fact that securitization rates tended to be low for jumbo mortgages.
[16] Distances are calculated using the latitude and longitude of counties as provided by the Census Bureau at http://www.census.gov/tiger/tms/gazetteer/county2k.txt. In areas comprised of multiple counties, the simple average of the county-specific latitudes and longitudes were calculated and used.

ensures that estimated median price for any area is anchored to a "true" median price. While this may not matter much in practice, it would seem preferable to link estimates of median prices to "true" median prices (rather than merely other estimates).

Column (4) in Table 1 shows that the nearby median provides some explanatory power. Estimating a 2.6 cent increase in an area's median price for every dollar increase in the nearby median price, the coefficient is significant at the five-percent error level.

As suggested by the fact that the association between the two median prices seemed to differ in 2007 and 2008 vis-à-vis earlier periods, the basic model can be improved further by allowing all of the coefficients to differ for each year. This approach is implemented here by simply breaking the sample into year-specific datasets. In doing so, the underlying regression becomes a basic cross-sectional model, with each observation representing a given geographic area. Columns 5a through 5g show statistics from and coefficients for the year-specific regressions.

The observation counts in the underlying regressions vary from 213 to 226 for 2002-2008. For 2008, the number of observations (areas) used in the model is 226, with about 125 representing MSAs, 60 representing μSAs and the remainder being unaffiliated counties. The basic composition of the modeling sample does not differ substantially in earlier periods.

The year-specific models shown in columns 5a-5g generally provide some additional explanatory power,[17] but the overall advantage is not particularly dramatic. The coefficient on the primary regressor—the median restricted price—varies from year to year, but is always statistically significant at the one-percent error level.

As alluded to previously, the year-specific differences in the statistical relationship stem in large part from variations in the relative contributions of the various datasets to the pooled data. For instance, observations from the LoanPerformance dataset, which include jumbo mortgages (with relatively high home values), comprise a large share of the pooled transactions in 2006, but are trivially represented in 2008.[18] Conversely, the proportion of FHA mortgages grows dramatically over the time period.[19]

Error Distribution

Figure 1 shows the distribution of relative errors arising from the year-specific models. Errors are expressed as a percentage of the unrestricted median price. Negative values arise when the actual median (the "unrestricted median") is less than the model's predicted value.

[17] F-Tests confirm the value of estimating year-specific models.

[18] For 2006 loans, the average home price associated with the LoanPerformance mortgages was about 45 percent higher than for Enterprise loans and about 155 percent above the FHA loans.

[19] It should be noted that the effect of the share variations on the coefficient is partially offset by changes in the divergence in home values across the various financing types. The average home value in the Enterprise dataset was 75 percent higher than the average in the FHA dataset in 2006, for example, but only 52 percent higher in 2008.

The bulk of the error distribution is clustered around zero, with a slight skewness toward negative errors. The cause of the predilection toward overestimation is unknown, but based on the relatively modest median error of +0.12 percent, the skewness is not particularly alarming.

Figure 2 plots the average percentage error against median home prices. The graph broadly shows that relative errors tend to be negative and large at the lowest end of the price spectrum; the unrestricted median price for DataQuick's county-recorder transactions appears extraordinarily low in some cases.

Because the primary purpose behind estimating local median prices in this exercise is to set loan limits *in high cost areas*, the relatively poor predictive performance at the lowest end of the price spectrum is not particularly troubling. Because a floor is set for loan limits in inexpensive areas (effectively setting loan limits at the same level in moderately-priced and inexpensive locations), there is no real impact of any bias for inexpensive areas. Also, importantly, the areas requiring median price estimates will be predominantly rural locales, where prices are low. Figure 2 plots a vertical line showing the median price (as measured in the financing-restricted data) in areas where the recorder-based data sample has no coverage.[20] While the $73,850 median value likely understates the "true" median in uncovered areas, it still reveals that non-overlap areas (areas requiring median price estimates) tend to be low priced.

For areas with somewhat higher median prices, relative errors tend to shrink in absolute terms. Estimation errors in the middle part of the price spectrum, including locales with home values between $150,000 and about $250,000, seem not to have a clear positive or negative bias. For the higher-priced areas, however, prediction errors tend to be positive and seem larger than the (absolute) errors in the middle-tier. Because positive errors reflect situations where the model underpredicts true median values, assuming that the underprediction phenomenon holds for high-cost areas *that require median estimates*, the bias could have significant loan limit implications. Further research will be undertaken to determine if model specification improvements can eliminate the bias.

Table 2 reports model-estimated and actual (recorder-based) median prices for the largest 25 metropolitan areas in the United States for 2008. The table, which also shows relative errors as a percentage of the recorder-based medians, generally indicates that predicted medians closely resemble actual medians in the latest full year. The predicted median was within 10 percent of the actual median in 19 out of the 23 metropolitan areas for which recorder-based medians were available.[21] The four remaining areas had errors of less than 20 percent.

Performance of FHFA and FHA/HUD Methodologies for Counties

With the basic estimation approach now described, the estimation accuracy of the new extrapolation methodology is compared against the accuracy of the median-estimation

[20] Note that the $73,800 median is the unweighted average median price across all nonoverlap areas for the years 2002-2008.

[21] Nondisclosure laws prohibit the calculation of recorder-based medians in the Dallas and Houston metropolitan areas.

methodology used by HUD. The analysis is done at the county-level using a dataset of actual and HUD-estimated median prices. The HUD data show county-level median prices for the period between January 1, 2008 and August 31, 2008,[22] where the reported "actual" median prices were determined by RadarLogic using its own county-recorder data. As indicated earlier, the HUD estimates were predominantly a function of self-reported home values from the American Community Survey data.

Figure 3 compares the distribution of errors arising from the HUD approach against those derived from the FHFA methodology. To ensure consistency, the FHFA medians and errors were calculated using data from the same time interval: January 1, 2008 – August 31, 2008. As is evident in the figure, the primary difference between the two distributions is that the HUD approach shows a much greater tendency to overpredict actual medians. Approximately 75 percent of HUD's estimation errors are less than zero (overpredictions) as compared with about 50 percent of FHFA's. More than half of HUD's predictions have errors of less than -10 percent, significantly more than the 20 percent share registered by the FHFA approach. HUD's tendency toward overprediction is not surprising given that its estimates ultimately rely on owner assessments of home values. Prior literature has shown a propensity of homeowners to overestimate the value of their homes.

Table 3 provides a more detailed analysis of the predictive accuracy of the two approaches. The median and mean errors, median and mean *absolute* errors and the root mean squared error are compared across the two methodologies for different sets of geographic areas.[23] The underlying errors, it should be noted, are calculated by comparing predicted values against different value for the "true" (recorder-based) median; HUD errors are calculated as the difference between the predicted median and the RadarLogic-based median price whereas the FHFA errors are the difference between the predicted values and the *DataQuick*-based median price.[24]

Columns (1) and (2) of Table 3 report the sample sizes and error statistics for all counties for which errors can be calculated in the respective datasets; Error estimates were available for more than 1,000 counties in the Radar Logic-based data and approximately 443 counties in the DataQuick data. The median and mean errors from the HUD methodology were -10.25 percent and -12.57 percent respectively, indicating a significant bias toward overprediction. By contrast, the FHFA approach produces average and median errors that tend to be relatively close to zero. The lack of substantial bias in fact arises by construction; for the overlap areas, in fitting the data, the regression will produce an average error close to zero. The average percentage error will then tend to be close to zero.

When the absolute errors and root mean squared errors are compared across the methods, columns (1) and (2) also indicate better performance (smaller errors) for the FHFA approach.

[22] This interval was the basis for determination of 2009 loan limits.

[23] The root mean squared error is a commonly used metric for summarizing predictive accuracy. Its value grows more rapidly (than the average absolute error) when very large errors are present.

[24] While the FHFA approach has not yet been calibrated on the RadarLogic-based median values (to enable error comparisons based on the same medians), empirical results indicate that the FHA's errors are very large when the DataQuick-based medians are used as the "true" medians.

At 20.89 percent, for example, the root mean squared error for the FHFA methodology, for example, was significantly less than the 26.38 percent value for HUD's ACS-based approach.

Columns (3) and (4) provide a better comparison of the relative accuracy of the two approaches by restricting the two data samples to the same set of counties. Errors are compared for the 401 counties for which both FHA and FHFA have computable errors. As with the prior results, the HUD approach is significantly biased (with a mean error of -9.1 percent) while the FHFA methodology is effectively unbiased. The mean and median absolute errors also show improved performance for the FHFA methodology. It does appear that the FHFA approach is plagued by some relatively large outlier errors, however; despite its smaller mean absolute error, FHFA approach produces a slightly larger root mean *squared* error.

Columns (1) – (4) report and compare prediction errors for individual county median price estimates. A more relevant issue might be how the respective methodologies perform in estimating local median prices under the high-cost county rule. That is—in the event that HUD or FHFA wish to (or are required to) continue setting metropolitan and micropolitan medians equal to the median price of the highest-priced component county, it would be useful to know whether the FHFA methodology exhibits estimation advantages.

Columns (5) and (6) report the models' success at estimating the medians under the high-cost county rule. The results, which are shown for 177 areas where both the HUD and FHA samples had complete geographic coverage, there were significant advantages for the FHFA methodology. As with the county-based estimates, the HUD predictions tended to be too high (without significant bias for FHFA);[25] median and mean errors for HUD's estimates were -9.3 and -9.5 percent respectively. Also, the associated mean absolute error and root mean squared error were larger than for the FHA predictions.

Loan Limit Implications of Using the FHFA Estimation Approach: High-Cost Areas

As indicated previously, for local loan limit determination, the lack of publicly available data for the majority of U.S. counties is not particularly pernicious. Median price data are currently available, whether from recorder-based sources (e.g., Radar Logic or DataQuick) or from MLS-based sources (NAR), for most high-cost areas. Because those sources are available for most high-cost areas and because HUD's median estimates are used only where those sources were unavailable, the practical implications of switching to an alternative median-estimation approach are not dramatic.

Table 4 reports that only 31 counties would have had different 2009 Enterprise loan limits had the FHFA median-estimates been used rather than HUD's median price estimates.[26] The loan limit differences were small in some cases, but were relatively significant for Judith Basin County, Montana and Richmond, Virginia. For the former, the loan limit determined under

[25] In this case, an "overprediction" means that the estimated high-cost-county median was higher than the "true" (recorder-based) high-cost-county median.
[26] The limit differences reported in the table reference the 2009 limits as determined under HERA, which set limits equal to 115 percent of median prices up to $625,500. Limits for all areas could not be below $417,000.

this approach would have been $53,300 higher. By contrast, Richmond, Virginia's limit would have been nearly $119,000 lower.

Although the loan limit effects of the alternative median estimates appear small and localized for Enterprise limits, they are somewhat larger for FHA limits. The greater effect arises from the fact that the "floor" for 2009 FHA loan limits, $271,050, is much lower than the $417,000 floor for Enterprise limits. The lower floor has the effect of directly tying loan limits to median price estimates for a larger number of areas. Changes in median price estimates thus tend to influence loan limits in a greater number of areas. Under the FHFA methodology, 82 counties would have had different FHA limits (versus the 31 counties with different Enterprise limits). Where differences would have existed, the magnitude of the divergence was often nontrivial; in 43 of the 82 counties, the FHFA-based loan limits would be at least $50,000 above or below the limit determined under HUD's calculation.

National Average and Median Home Price: Estimation and Loan Limits

Although the described technique and the empirical results thus far have focused on estimating *local median* prices, the same basic approach can be applied to estimating local average prices. The median-orientation has been a function of the need to develop an approach that can be used for setting local loan limits for high-cost areas.

In setting loan limits, FHFA is also required to establish a way of gauging changes in the *national average* home price. FHFA's enabling legislation, the Housing and Economic Recovery Act of 2008 (HERA), ties changes in the national loan limit to changes in the national average home price. The FHFA Director, specifically, is required to: "…establish and maintain a method for assessing the national average 1-family house price for use for adjusting the conforming loan limitations of the enterprises…"[27] The Act does not specify whether it contemplates the use of the *arithmetic* or *geometric* average home price. The latter, which might be approximated with a median price metric, would effectively align the local and national bases for loan limits: both would be based on median prices.

One strategy for implementing this provision might be to estimate local average (or median) prices using the methodology described in this paper and then to "build up" a national measure as a weighted average of the local estimates. The national measure could be constructed as a weighted average of the local price aggregates using Census Bureau estimates for the number of housing units in the respective areas. This type of weighting-based approach would be attractive because it would minimize the effect of geographic shifts in relative transaction volumes. If not controlled for, geographic shifts in sales volumes over time can make changes in average and median values misleading measures of price movements.

Using the basic methodology described in this paper, a weighted national average home price has been estimated for the first eight months of 2008. This county-based, average price measure is reported in Table 6 and is compared with the average price reported by the

[27] See Section 1322 of the Act.

National Association of Realtors (NAR). The two figures are very similar, with the constructed weighted average price less than $4,000 below the NAR estimate.[28]

CONCLUSIONS AND NEXT STEPS

Although the basic approach toward estimating local median prices and average U.S. prices shows some promise for approximating house price statistics, work remains to be done. It may be useful, for example, to further investigate the causes of the instability of the statistical relationship between the restricted and unrestricted medians. While some explanations are available for the significant changes in the coefficient estimates, additional analysis is necessary to ensure that the instability (and lower explanatory power of the model) does not reflect data errors or model breakdown.

Given the methodology's nascency, FHFA intends to track its performance for several quarters. It also plans to analyze cross-sectional differences in the relative performance of the model; HUD's approach for estimating medians could prove to be consistently more precise in some geographic areas. This would open up the possibility for a hybrid model that would exploit the relative benefits of both approaches.

The author welcomes public comments about the methodology described in this paper. Table 7 shows median price estimates for 2008 that would be produced with the approach. Median price estimates are shown for every metropolitan area in the country.

[28] If the approach to estimating national average home prices is modified slightly so that recorder-based local average prices are used where available (instead of estimated mean prices), the resulting average price would be similar; the weighted average U.S. price would be $238,600 rather than $244,392.

Table 1: Models of Unrestricted Median Prices--Coefficient Estimates

	[1]	[2]	[3]	[4]	[5a]	[5b]	[5c]	[5d]	[5e]	[5f]	[5g]
Dependant Variable	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price	Median DQ Price
Year(s)	2002-2008	2002-2008	2002-2008	2002-2008	2002	2003	2004	2005	2006	2007	2008
Specification	1	2	3	4	5	5	5	5	5	5	5
Variable Name											
median_price_restricted	**1.11145*** **	**1.09157*** **	**1.22932*** **	**1.20539*** **	**1.13895*** **	**1.18359*** **	**1.11783*** **	**1.12375*** **	**1.28966*** **	**1.3711*** **	**1.66851*** **
std err	0.009	0.01	0.041	0.042	0.068	0.059	0.098	0.101	0.112	0.121	0.176
median_price_restricted07		**0.06247*** **	**0.06191*** **	**0.06007*** **							
std err		0.011	0.011	0.011							
median_price_above200			**-0.07783*** **	**-0.07757*** **	**-0.07976* **	**-0.07481** **	**-0.03665**	**-0.05159**	**-0.10109**	**-0.11212* **	**-0.18759** **
std err			0.024	0.024	0.041	0.035	0.054	0.057	0.062	0.066	0.091
median_price_above400			**-0.03598** **	**-0.03536** **	**-0.01398**	**-0.10603*** **	**-0.09672*** **	**-0.04139**	**-0.07007* **	**-0.03423**	**0.05684**
std err			0.015	0.015	0.035	0.026	0.036	0.034	0.038	0.041	0.071
median_price_nearby				**0.02592** **	**0.0582** **	**0.05769*** **	**0.07034** **	**0.0481** **	**-0.00463**	**0.01591**	**-0.02033**
std err				0.01	0.026	0.021	0.027	0.022	0.024	0.023	0.031
Observations	**1,552**	**1,552**	**1,552**	**1,552**	**213**	**218**	**221**	**223**	**226**	**225**	**226**
R-squared	**0.891**	**0.893**	**0.894**	**0.894**	**0.914**	**0.940**	**0.907**	**0.935**	**0.919**	**0.914**	**0.786**

Note: Interecept terms not show. Significance at 10 percent, 5 percent, and 1 percent error levels denoted by *, **, and *** respectively.

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Table 2: Median Price Predictions and Errors for 25 Most Populated Metropolitan Areas (2008 Data)

CBSA Number	Metro Name	Predicted Median House Value	DataQuick Median House Value	Percent Error
35620	New York-Northern New Jersey-Long Island, NY-NJ-PA Metro Area	$ 445,800	$ 395,000	-12.861%
31100	Los Angeles-Long Beach-Santa Ana, CA Metro Area	$ 450,632	$ 430,000	-4.798%
16980	Chicago-Naperville-Joliet, IL-IN-WI Metro Area	$ 214,625	$ 229,000	6.277%
19100	Dallas-Fort Worth-Arlington, TX Metro Area	$ 124,335	N/A	-
37980	Philadelphia-Camden-Wilmington, PA-NJ-DE-MD Metro Area	$ 204,745	$ 226,000	9.405%
26420	Houston-Sugar Land-Baytown, TX Metro Area	$ 129,771	N/A	-
33100	Miami-Fort Lauderdale-Pompano Beach, FL Metro Area	$ 243,845	$ 262,500	7.107%
12060	Atlanta-Sandy Springs-Marietta, GA Metro Area	$ 148,265	$ 139,000	-6.666%
47900	Washington-Arlington-Alexandria, DC-VA-MD-WV Metro Area	$ 349,986	$ 364,000	3.850%
14460	Boston-Cambridge-Quincy, MA-NH Metro Area	$ 364,100	$ 335,000	-8.686%
19820	Detroit-Warren-Livonia, MI Metro Area	$ 96,888	$ 89,900	-7.774%
38060	Phoenix-Mesa-Scottsdale, AZ Metro Area	$ 198,778	$ 205,000	3.035%
41860	San Francisco-Oakland-Fremont, CA Metro Area	$ 500,706	$ 465,000	-7.679%
40140	Riverside-San Bernardino-Ontario, CA Metro Area	$ 235,403	$ 240,000	1.915%
42660	Seattle-Tacoma-Bellevue, WA Metro Area	$ 381,666	$ 350,000	-9.047%
33460	Minneapolis-St. Paul-Bloomington, MN-WI Metro Area	$ 192,712	$ 205,000	5.994%
41740	San Diego-Carlsbad-San Marcos, CA Metro Area	$ 407,190	$ 390,000	-4.408%
41180	St. Louis, MO-IL Metro Area	$ 133,439	$ 135,000	1.156%
45300	Tampa-St. Petersburg-Clearwater, FL Metro Area	$ 151,428	$ 158,000	4.160%
12580	Baltimore-Towson, MD Metro Area	$ 262,575	$ 250,000	-5.030%
19740	Denver-Aurora, CO Metro Area	$ 183,743	$ 215,000	14.538%
38300	Pittsburgh, PA Metro Area	$ 88,211	$ 108,000	18.323%
38900	Portland-Vancouver-Beaverton, OR-WA Metro Area	$ 270,777	$ 268,000	-1.036%
17140	Cincinnati-Middletown, OH-KY-IN Metro Area	$ 114,902	$ 133,450	13.899%
40900	Sacramento--Arden-Arcade--Roseville, CA Metro Area	$ 254,306	$ 250,000	-1.723%

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Table 3: Estimation Errors HUD Method vs. FHFA Approach

	[1]	[2]	[3]	[4]	[5]	[6]
	All Counties available in respective datasets		**Only Counties where both RadarLogic and DataQuick have Coverage**		**"Areas" (MSA/µSAs/counties) where both RadarLogic and FHFA's Recorder-Data (from DataQuick) have full Geographic Coverage** (Uses High-Cost-County Rule for Aggregation)	
Estimation Model	HUD/FHA	FHFA	HUD/FHA	FHFA	HUD/FHA	FHFA
"True" Median	RadarLogic	DataQuick	RadarLogic	DataQuick	RadarLogic	DataQuick
# of Locations	**1,067**	**443**	**401**	**401**	**177**	**177**
Median Error	-10.25%	-0.31%	-8.53%	0.22%	-9.25%	-1.05%
Median Absolute Error	12.80%	8.69%	10.36%	8.79%	11.84%	9.07%
Mean Error	-12.57%	-2.36%	-9.05%	-2.00%	-9.50%	-0.76%
Mean Absolute Error	18.21%	13.36%	13.81%	13.17%	14.79%	12.91%
Root Mean Squared Error	26.38%	20.89%	18.64%	20.64%	19.72%	18.48%

Note: Relative errors calculated as (Actual Price-Estimated Price)/Actual Price

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, RadarLogic, and First American CoreLogic ("LoanPerformance Data").

Table 4: Areas with Different 2009 Enterprise Loan Limits under the FHFA Median Price Estimates

FIPS Number	County Name	State	Metro Name	Loan Limit Utilizing ACS Median	Loan Limit Utitilizing FHFA Median	Difference
51570	COLONIAL HEIGHT	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51149	PRINCE GEORGE	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51036	CHARLES CITY	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51007	AMELIA	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51085	HANOVER	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51053	DINWIDDIE	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51730	PETERSBURG	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51109	LOUISA	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51101	KING WILLIAM	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51049	CUMBERLAND	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51033	CAROLINE	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51041	CHESTERFIELD	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51097	KING AND QUEEN	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51670	HOPEWELL	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51075	GOOCHLAND	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51760	RICHMOND IND	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51145	POWHATAN	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51183	SUSSEX	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51127	NEW KENT	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
51087	HENRICO	VA	RICHMOND, VA (MSA)	$535,900	$417,000	-$118,900
69120	TINIAN	MP	NON-METRO	$532,450	$417,000	-$115,450
69110	SAIPAN	MP	NON-METRO	$529,000	$417,000	-$112,000
69085	NORTHERN ISLAND	MP	NON-METRO	$524,400	$417,000	-$107,400
51103	LANCASTER	VA	NON-METRO	$442,750	$417,000	-$25,750
51125	NELSON	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$417,000	-$20,000
51079	GREENE	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$417,000	-$20,000
51540	CHARLOTTESVILLE	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$417,000	-$20,000
51003	ALBEMARLE	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$417,000	-$20,000
51065	FLUVANNA	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$417,000	-$20,000
08053	HINSDALE	CO	NON-METRO	$423,200	$417,000	-$6,200
30045	JUDITH BASIN	MT	NON-METRO	$417,000	$470,350	$53,350

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Table 5: Areas with Different 2009 FHA Loan Limits under the FHFA Median Price Estimates

County Name	State	Metro Name	FHA Limit Utilizing ACS Median	FHA Limit Utitilizing FHFA Median	Difference
COLONIAL HEIGHT	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
PRINCE GEORGE	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
CHARLES CITY	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
AMELIA	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
HANOVER	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
DINWIDDIE	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
PETERSBURG	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
LOUISA	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
KING WILLIAM	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
CUMBERLAND	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
CAROLINE	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
CHESTERFIELD	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
KING AND QUEEN	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
HOPEWELL	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
GOOCHLAND	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
RICHMOND IND	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
POWHATAN	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
SUSSEX	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
NEW KENT	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
HENRICO	VA	RICHMOND, VA (MSA)	$535,900	$317,400	-$218,500
HINSDALE	CO	NON-METRO	$423,200	$271,050	-$152,150
YAKUTAT CITY	AK	NON-METRO	$381,800	$271,050	-$110,750
NELSON	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$332,350	-$104,650
GREENE	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$332,350	-$104,650
CHARLOTTESVILLE	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$332,350	-$104,650
ALBEMARLE	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$332,350	-$104,650
FLUVANNA	VA	CHARLOTTESVILLE, VA (MSA)	$437,000	$332,350	-$104,650
LITCHFIELD	CT	TORRINGTON, CT (MICRO)	$357,650	$271,050	-$86,600
LOS ALAMOS	NM	LOS ALAMOS, NM (MICRO)	$380,650	$300,150	-$80,500
ALEUTIANS WEST	AK	NON-METRO	$349,600	$271,050	-$78,550
WEBER	UT	OGDEN-CLEARFIELD, UT (MSA)	$389,850	$312,800	-$77,050
MORGAN	UT	OGDEN-CLEARFIELD, UT (MSA)	$389,850	$312,800	-$77,050
DAVIS	UT	OGDEN-CLEARFIELD, UT (MSA)	$389,850	$312,800	-$77,050
ST. CROIX	VI	NON-METRO	$327,750	$271,050	-$56,700
LANCASTER	VA	NON-METRO	$442,750	$391,000	-$51,750
SANTA FE	NM	SANTA FE, NM (MSA)	$368,000	$322,000	-$46,000
GALLATIN	MT	BOZEMAN, MT (MICRO)	$346,150	$304,750	-$41,400
KETCHIKAN GATEW	AK	KETCHIKAN, AK (MICRO)	$322,000	$281,750	-$40,250
FLATHEAD	MT	KALISPELL, MT (MICRO)	$301,300	$271,050	-$30,250
WRANGELL-PETERS	AK	NON-METRO	$301,300	$271,050	-$30,250

Table 5: Areas with Different 2009 FHA Loan Limits under the FHFA Median Price Estimates

County Name	State	Metro Name	FHA Limit Utilizing ACS Median	FHA Limit Utitilizing FHFA Median	Difference
NORTH SLOPE	AK	NON-METRO	$301,300	$271,050	-$30,250
CLALLAM	WA	PORT ANGELES, WA (MICRO)	$296,700	$271,050	-$25,650
MISSOULA	MT	MISSOULA, MT (MSA)	$282,900	$271,050	-$11,850
COOK	MN	NON-METRO	$282,900	$271,050	-$11,850
SWEET GRASS	MT	NON-METRO	$280,600	$271,050	-$9,550
KNOX	ME	ROCKLAND, ME (MICRO)	$279,450	$271,050	-$8,400
WASHINGTON	UT	ST. GEORGE, UT (MSA)	$278,300	$271,050	-$7,250
BENNINGTON	VT	BENNINGTON, VT (MICRO)	$277,150	$271,050	-$6,100
HARRISONBURG	VA	HARRISONBURG, VA (MSA)	$277,150	$271,050	-$6,100
ROCKINGHAM	VA	HARRISONBURG, VA (MSA)	$277,150	$271,050	-$6,100
JEFFERSON	MT	HELENA, MT (MICRO)	$277,150	$271,050	-$6,100
LEWIS AND CLARK	MT	HELENA, MT (MICRO)	$277,150	$271,050	-$6,100
WINDHAM	CT	WILLIMANTIC, CT (MICRO)	$271,400	$271,050	-$350
HANCOCK	ME	NON-METRO	$271,400	$271,050	-$350
CROOK	OR	PRINEVILLE, OR (MICRO)	$271,050	$271,400	$350
HIGHLAND	VA	NON-METRO	$271,050	$272,550	$1,500
BLANCO	TX	NON-METRO	$271,050	$273,700	$2,650
GRAND	UT	NON-METRO	$271,050	$273,700	$2,650
DILLINGHAM	AK	NON-METRO	$271,050	$274,850	$3,800
KODIAK ISLAND	AK	KODIAK, AK (MICRO)	$317,400	$323,150	$5,750
WILKINSON	MS	NON-METRO	$271,050	$280,600	$9,550
CURRY	OR	BROOKINGS, OR (MICRO)	$327,750	$342,700	$14,950
NORTHUMBERLAND	VA	NON-METRO	$318,550	$333,500	$14,950
BONNER	ID	NON-METRO	$271,050	$290,950	$19,900
SUBLETTE	WY	NON-METRO	$276,000	$310,500	$34,500
ADAMS	ID	NON-METRO	$271,050	$308,200	$37,150
BAKER	GA	ALBANY, GA (MSA)	$271,050	$309,350	$38,300
LEE	GA	ALBANY, GA (MSA)	$271,050	$309,350	$38,300
WORTH	GA	ALBANY, GA (MSA)	$271,050	$309,350	$38,300
TERRELL	GA	ALBANY, GA (MSA)	$271,050	$309,350	$38,300
DOUGHERTY	GA	ALBANY, GA (MSA)	$271,050	$309,350	$38,300
ANCHORAGE	AK	ANCHORAGE, AK (MSA)	$290,950	$333,500	$42,550
MATANUSKA-SUSIT	AK	ANCHORAGE, AK (MSA)	$290,950	$333,500	$42,550
MIDDLESEX	VA	NON-METRO	$271,050	$317,400	$46,350
REAL	TX	NON-METRO	$271,050	$322,000	$50,950
MENOMINEE	WI	NON-METRO	$271,050	$325,450	$54,400
TAOS	NM	TAOS, NM (MICRO)	$271,050	$336,950	$65,900
WASATCH	UT	HEBER, UT (MICRO)	$325,450	$401,350	$75,900
SITKA	AK	NON-METRO	$340,400	$427,800	$87,400
WINSTON	AL	NON-METRO	$271,050	$401,350	$130,300

Table 5: Areas with Different 2009 FHA Loan Limits under the FHFA Median Price Estimates

County Name	State	Metro Name	FHA Limit Utilizing ACS Median	FHA Limit Utitilizing FHFA Median	Difference
RICH	UT	NON-METRO	$271,050	$411,700	$140,650
JUDITH BASIN	MT	NON-METRO	$271,050	$470,350	$199,300

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Table 6: U.S. Home Price Estimates

Mean Prices for January 1, 2008 - August 31, 2008

Source	Average Price
FHFA--Weighted Average of Estimated Average Prices for Counties	$244,392
National Association of Realtors (NAR)	$247,663

Notes:

FHFA estimates are constructed as the weighted average of county-level estimates, where the weights are the relative shares of one-unit properties in the respective counties as reported in the 2000 Census.

The NAR estimate is the average of the monthly average home prices reported for the first eight months of 2008.

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Abilene, TX	$76,560
Akron, OH	$108,119
Albany, GA	$101,719
Albany-Schenectady-Troy, NY	$188,972
Albuquerque, NM	$177,460
Alexandria, LA	$146,330
Allentown-Bethlehem-Easton, PA	$213,888
Altoona, PA	$55,176
Amarillo, TX	$87,572
Ames, IA	$147,440
Anchorage, AK	$270,393
Anderson, IN	$65,574
Anderson, SC	$89,379
Ann Arbor, MI	$194,558
Anniston-Oxford, AL	$89,278
Appleton, WI	$148,164
Asheville, NC	$190,276
Athens-Clarke County, GA	$126,313
Atlanta-Sandy Springs-Marietta, GA	$148,265
Atlantic City-Hammonton, NJ	$235,672
Auburn-Opelika, AL	$142,935
Augusta-Richmond County, GA	$94,383
Austin-Round Rock, TX	$165,184
Bakersfield, CA	$213,420
Baltimore-Towson, MD	$262,575
Bangor, ME	$126,505
Barnstable Town, MA	$307,699
Baton Rouge, LA	$177,739
Battle Creek, MI	$39,952
Bay City, MI	$41,889
Beaumont-Port Arthur, TX	$92,645
Bellingham, WA	$285,492
Bend, OR	$249,447
Billings, MT	$184,405
Binghamton, NY	$61,961
Birmingham-Hoover, AL	$128,011
Bismarck, ND	$160,719
Blacksburg-Christiansburg-Radford, VA	$180,944
Bloomington, IN	$105,943

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Bloomington-Normal, IL	$151,827
Boise City-Nampa, ID	$164,158
Boston-Cambridge-Quincy, MA-NH Metropolitan Stati	$364,100
Boulder, CO	$271,953
Bowling Green, KY	$112,382
Bradenton-Sarasota-Venice, FL	$156,915
Bremerton-Silverdale, WA	$255,606
Bridgeport-Stamford-Norwalk, CT	$441,462
Brownsville-Harlingen, TX	$50,415
Brunswick, GA	$141,623
Buffalo-Niagara Falls, NY	$85,169
Burlington, NC	$114,560
Burlington-South Burlington, VT	$250,166
Canton-Massillon, OH	$64,850
Cape Coral-Fort Myers, FL	$117,541
Carson City, NV	$222,922
Casper, WY	$187,742
Cedar Rapids, IA	$115,302
Champaign-Urbana, IL	$131,987
Charleston, WV	$112,349
Charleston-North Charleston-Summerville, SC	$190,283
Charlotte-Gastonia-Concord, NC	$143,890
Charlottesville, VA	$260,750
Chattanooga, TN	$114,895
Cheyenne, WY	$163,690
Chicago-Naperville-Joliet, IL-IN-WI Metropolitan	$216,053
Chico, CA	$228,314
Cincinnati-Middletown, OH	$115,507
Clarksville, TN	$98,950
Cleveland, TN	$112,226
Cleveland-Elyria-Mentor, OH	$98,220
Coeur d'Alene, ID	$205,442
College Station-Bryan, TX	$117,632
Colorado Springs, CO	$188,537
Columbia, MO	$114,487
Columbia, SC	$117,405
Columbus, GA	$118,477
Columbus, IN	$90,479
Columbus, OH	$140,514

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Corpus Christi, TX	$105,993
Corvallis, OR	$249,732
Cumberland, MD	$76,866
Dallas-Fort Worth-Arlington, TX Metropolitan Stat	$120,552
Dalton, GA	$75,518
Danville, IL	$7,500
Danville, VA	$78,687
Davenport-Moline-Rock Island, IA	$81,586
Dayton, OH	$75,463
Decatur, AL	$77,333
Decatur, IL	$40,871
Deltona-Daytona Beach-Ormond Beach, FL	$140,408
Denver-Aurora, CO	$174,783
Des Moines-West Des Moines, IA	$149,874
Detroit-Warren-Livonia, MI Metropolitan Statistic	$96,888
Dothan, AL	$93,509
Dover, DE	$191,237
Dubuque, IA	$101,036
Duluth, MN	$118,477
Durham, NC	$185,944
Eau Claire, WI	$117,296
El Centro, CA	$206,694
Elizabethtown, KY	$90,071
Elkhart-Goshen, IN	$98,350
Elmira, NY	$29,822
El Paso, TX	$104,397
Erie, PA	$71,412
Eugene-Springfield, OR	$214,276
Evansville, IN	$76,452
Fairbanks, AK	$199,309
Fargo, ND	$129,017
Farmington, NM	$191,854
Fayetteville, NC	$97,092
Fayetteville-Springdale-Rogers, AR	$128,539
Flagstaff, AZ	$301,252
Flint, MI	$15,026
Florence, SC	$114,629
Florence-Muscle Shoals, AL	$77,688
Fond du Lac, WI	$91,434

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Fort Collins-Loveland, CO	$175,672
Fort Smith, AR	$73,015
Fort Walton Beach-Crestview-Destin, FL	$181,110
Fort Wayne, IN	$53,633
Fresno, CA	$208,868
Gadsden, AL	$80,206
Gainesville, FL	$164,875
Gainesville, GA	$128,071
Glens Falls, NY	$138,703
Goldsboro, NC	$98,251
Grand Forks, ND	$123,177
Grand Junction, CO	$187,505
Grand Rapids-Wyoming, MI	$80,203
Great Falls, MT	$126,939
Greeley, CO	$129,344
Green Bay, WI	$104,782
Greensboro-High Point, NC	$114,275
Greenville, NC	$125,002
Greenville-Mauldin-Easley, SC	$129,159
Gulfport-Biloxi, MS	$127,684
Hagerstown-Martinsburg, MD	$208,915
Hanford-Corcoran, CA	$212,681
Harrisburg-Carlisle, PA	$158,198
Harrisonburg, VA	$207,080
Hartford-West Hartford-East Hartford, CT	$234,301
Hattiesburg, MS	$156,049
Hickory-Lenoir-Morganton, NC	$104,584
Hinesville-Fort Stewart, GA	$93,653
Holland-Grand Haven, MI	$118,664
Honolulu, HI	$795,420
Hot Springs, AR	$107,203
Houma-Bayou Cane-Thibodaux, LA	$156,549
Houston-Sugar Land-Baytown, TX	$121,345
Huntington-Ashland, WV	$66,286
Huntsville, AL	$145,530
Idaho Falls, ID	$150,534
Indianapolis-Carmel, IN	$94,263
Iowa City, IA	$187,048
Ithaca, NY	$178,737

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Jackson, MI	$46,082
Jackson, MS	$131,814
Jackson, TN	$85,054
Jacksonville, FL	$164,747
Jacksonville, NC	$136,323
Janesville, WI	$123,299
Jefferson City, MO	$88,753
Johnson City, TN	$121,943
Johnstown, PA	$25,142
Jonesboro, AR	$81,175
Joplin, MO	$50,210
Kalamazoo-Portage, MI	$94,512
Kankakee-Bradley, IL	$138,479
Kansas City, MO	$120,455
Kennewick-Pasco-Richland, WA	$145,971
Killeen-Temple-Fort Hood, TX	$88,433
Kingsport-Bristol-Bristol, TN	$89,558
Kingston, NY	$243,316
Knoxville, TN	$148,344
Kokomo, IN	$20,263
La Crosse, WI	$129,506
Lafayette, IN	$74,572
Lafayette, LA	$161,889
Lake Charles, LA	$124,540
Lake Havasu City-Kingman, AZ	$171,057
Lakeland-Winter Haven, FL	$136,198
Lancaster, PA	$177,386
Lansing-East Lansing, MI	$79,997
Laredo, TX	$96,888
Las Cruces, NM	$158,807
Las Vegas-Paradise, NV	$191,077
Lawrence, KS	$165,504
Lawton, OK	$61,829
Lebanon, PA	$122,954
Lewiston, ID	$146,777
Lewiston-Auburn, ME	$143,691
Lexington-Fayette, KY	$123,989
Lima, OH	$37,921
Lincoln, NE	$97,315

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Little Rock-North Little Rock-Conway, AR	$111,792
Logan, UT	$161,045
Longview, TX	$106,786
Longview, WA	$187,137
Los Angeles-Long Beach-Santa Ana, CA Metropolitan	$450,632
Louisville/Jefferson County, KY	$124,823
Lubbock, TX	$68,384
Lynchburg, VA	$159,379
Macon, GA	$105,963
Madera, CA	$203,563
Madison, WI	$223,194
Manchester-Nashua, NH	$259,695
Mansfield, OH	$16,463
McAllen-Edinburg-Mission, TX	$80,966
Medford, OR	$226,648
Memphis, TN	$124,072
Merced, CA	$156,650
Miami-Fort Lauderdale-Pompano Beach, FL Metropoli	$243,845
Michigan City-La Porte, IN	$93,678
Midland, TX	$152,286
Milwaukee-Waukesha-West Allis, WI	$210,733
Minneapolis-St. Paul-Bloomington, MN	$197,628
Missoula, MT	$213,290
Mobile, AL	$112,667
Modesto, CA	$178,518
Monroe, LA	$123,471
Monroe, MI	$101,894
Montgomery, AL	$103,520
Morgantown, WV	$166,132
Morristown, TN	$109,162
Mount Vernon-Anacortes, WA	$272,615
Muncie, IN	$29,972
Muskegon-Norton Shores, MI	$28,106
Myrtle Beach-North Myrtle Beach-Conway, SC	$179,159
Napa, CA	$537,612
Naples-Marco Island, FL	$214,226
Nashville-Davidson--Murfreesboro--Franklin, TN	$148,748
New Haven-Milford, CT	$243,605
New Orleans-Metairie-Kenner, LA	$181,076

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
New York-Northern New Jersey-Long Island, NY-NJ-P	$511,544
Niles-Benton Harbor, MI	$121,208
Norwich-New London, CT	$238,310
Ocala, FL	$173,090
Ocean City, NJ	$315,736
Odessa, TX	$78,454
Ogden-Clearfield, UT	$208,306
Oklahoma City, OK	$92,604
Olympia, WA	$251,863
Omaha-Council Bluffs, NE	$114,220
Orlando-Kissimmee, FL	$205,748
Oshkosh-Neenah, WI	$100,957
Owensboro, KY	$84,017
Oxnard-Thousand Oaks-Ventura, CA	$469,098
Palm Bay-Melbourne-Titusville, FL	$127,227
Palm Coast, FL	$151,399
Panama City-Lynn Haven, FL	$177,773
Parkersburg-Marietta-Vienna, WV	$63,518
Pascagoula, MS	$111,933
Pensacola-Ferry Pass-Brent, FL	$113,309
Peoria, IL	$88,313
Philadelphia-Camden-Wilington, PA-NJ-DE-MD Metrop	$204,667
Phoenix-Mesa-Scottsdale, AZ	$198,778
Pine Bluff, AR	$23,361
Pittsburgh, PA	$90,214
Pittsfield, MA	$178,679
Pocatello, ID	$95,807
Portland-South Portland-Biddeford, ME	$224,746
Portland-Vancouver-Beaverton, OR	$279,276
Port St. Lucie, FL	$141,921
Poughkeepsie-Newburgh-Middletown, NY	$306,609
Prescott, AZ	$184,883
Providence-New Bedford-Fall River, RI	$235,200
Provo-Orem, UT	$220,562
Pueblo, CO	$85,003
Punta Gorda, FL	$140,230
Racine, WI	$165,011
Raleigh-Cary, NC	$179,173
Rapid City, SD	$147,088

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Reading, PA	$147,147
Redding, CA	$225,777
Reno-Sparks, NV	$237,023
Richmond, VA	$190,063
Riverside-San Bernardino-Ontario, CA	$235,403
Roanoke, VA	$150,828
Rochester, MN	$161,208
Rochester, NY	$94,262
Rockford, IL	$104,312
Rocky Mount, NC	$110,765
Rome, GA	$81,603
Sacramento--Arden-Arcade--Roseville, CA	$254,306
Saginaw-Saginaw Township North, MI	$31,878
St. Cloud, MN	$147,356
St. George, UT	$220,696
St. Joseph, MO	$53,934
St. Louis, MO	$130,466
Salem, OR	$176,143
Salinas, CA	$334,833
Salisbury, MD	$188,341
Salt Lake City, UT	$224,862
San Angelo, TX	$73,365
San Antonio, TX	$118,428
San Diego-Carlsbad-San Marcos, CA	$407,190
Sandusky, OH	$51,390
San Francisco-Oakland-Fremont, CA Metropolitan St	$500,706
San Jose-Sunnyvale-Santa Clara, CA	$641,238
San Luis Obispo-Paso Robles, CA	$449,995
Santa Barbara-Santa Maria-Goleta, CA	$311,426
Santa Cruz-Watsonville, CA	$770,524
Santa Fe, NM	$305,973
Santa Rosa-Petaluma, CA	$409,738
Savannah, GA	$150,708
Scranton--Wilkes-Barre, PA	$91,431
Seattle-Tacoma-Bellevue, WA	$381,666
Sebastian-Vero Beach, FL	$139,502
Sheboygan, WI	$114,793
Sherman-Denison, TX	$52,727
Shreveport-Bossier City, LA	$115,129

Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Sioux City, IA	$47,260
Sioux Falls, SD	$140,530
South Bend-Mishawaka, IN	$40,285
Spartanburg, SC	$92,278
Spokane, WA	$180,759
Springfield, IL	$76,952
Springfield, MA	$197,101
Springfield, MO	$82,079
Springfield, OH	$41,491
State College, PA	$192,939
Stockton, CA	$205,480
Sumter, SC	$92,714
Syracuse, NY	$77,541
Tallahassee, FL	$140,486
Tampa-St. Petersburg-Clearwater, FL	$151,428
Terre Haute, IN	$22,517
Texarkana, TX	$78,755
Toledo, OH	$81,872
Topeka, KS	$64,331
Trenton-Ewing, NJ	$242,523
Tucson, AZ	$179,052
Tulsa, OK	$97,095
Tuscaloosa, AL	$127,262
Tyler, TX	$109,956
Utica-Rome, NY	$37,972
Valdosta, GA	$115,594
Vallejo-Fairfield, CA	$296,759
Victoria, TX	$90,101
Vineland-Millville-Bridgeton, NJ	$179,112
Virginia Beach-Norfolk-Newport News, VA	$200,363
Visalia-Porterville, CA	$201,635
Waco, TX	$89,684
Warner Robins, GA	$105,963
Washington-Arlington-Alexandria, DC-VA-MD-WV Metr	$352,863
Waterloo-Cedar Falls, IA	$85,788
Wausau, WI	$96,440
Weirton-Steubenville, WV	$6,452
Wenatchee, WA	$231,845
Wheeling, WV	$39,325

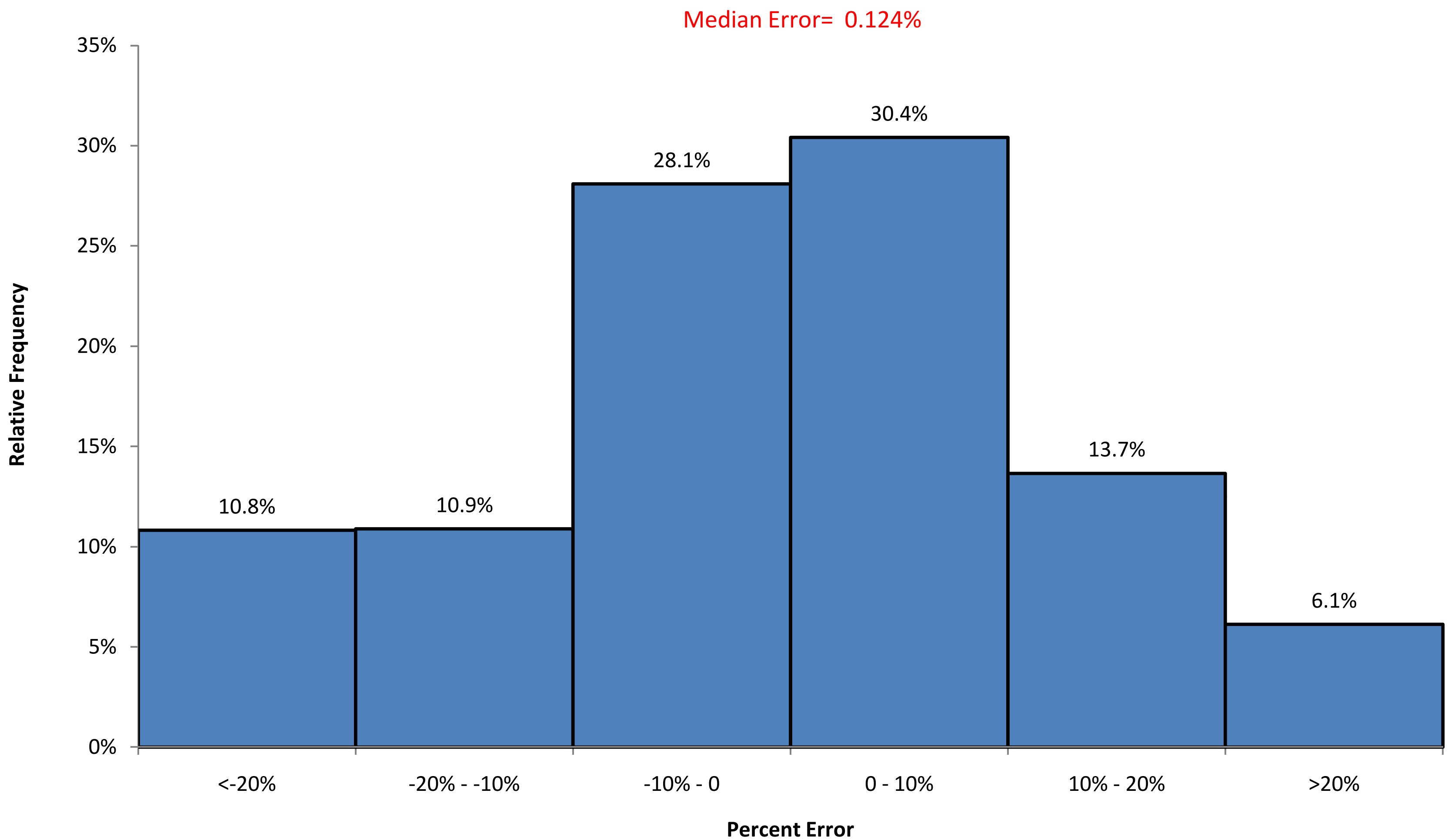
Table 7: Estimated Median House Price in 2008 based on FHFA Methodology

Metro Name	Median House Price (Estimated)
Wichita, KS	$68,269
Wichita Falls, TX	$44,643
Williamsport, PA	$83,256
Wilmington, NC	$204,753
Winchester, VA	$187,875
Winston-Salem, NC	$124,035
Worcester, MA	$235,200
Yakima, WA	$151,728
York-Hanover, PA	$160,495
Youngstown-Warren-Boardman, OH	$31,480
Yuba City, CA	$188,275
Yuma, AZ	$180,212

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Figure 1: Relative Frequency of Prediction Error (as % of Full-Market Median)

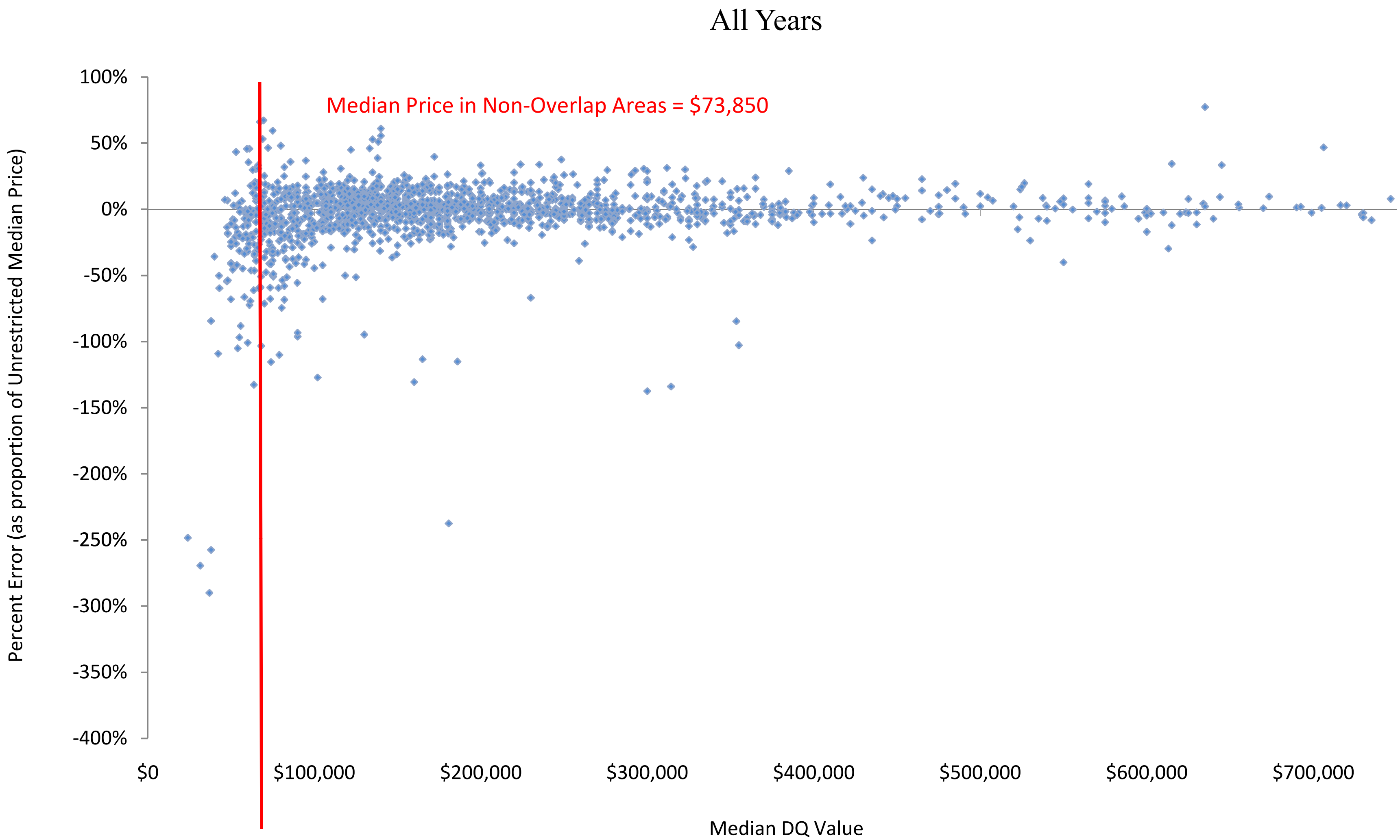
All Years



Note: Full-market median estimated from county-recorder data supplied by DataQuick Information Systems. Prediction error calculated as (Actual Price-Estimated Price)/Actual Price.

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

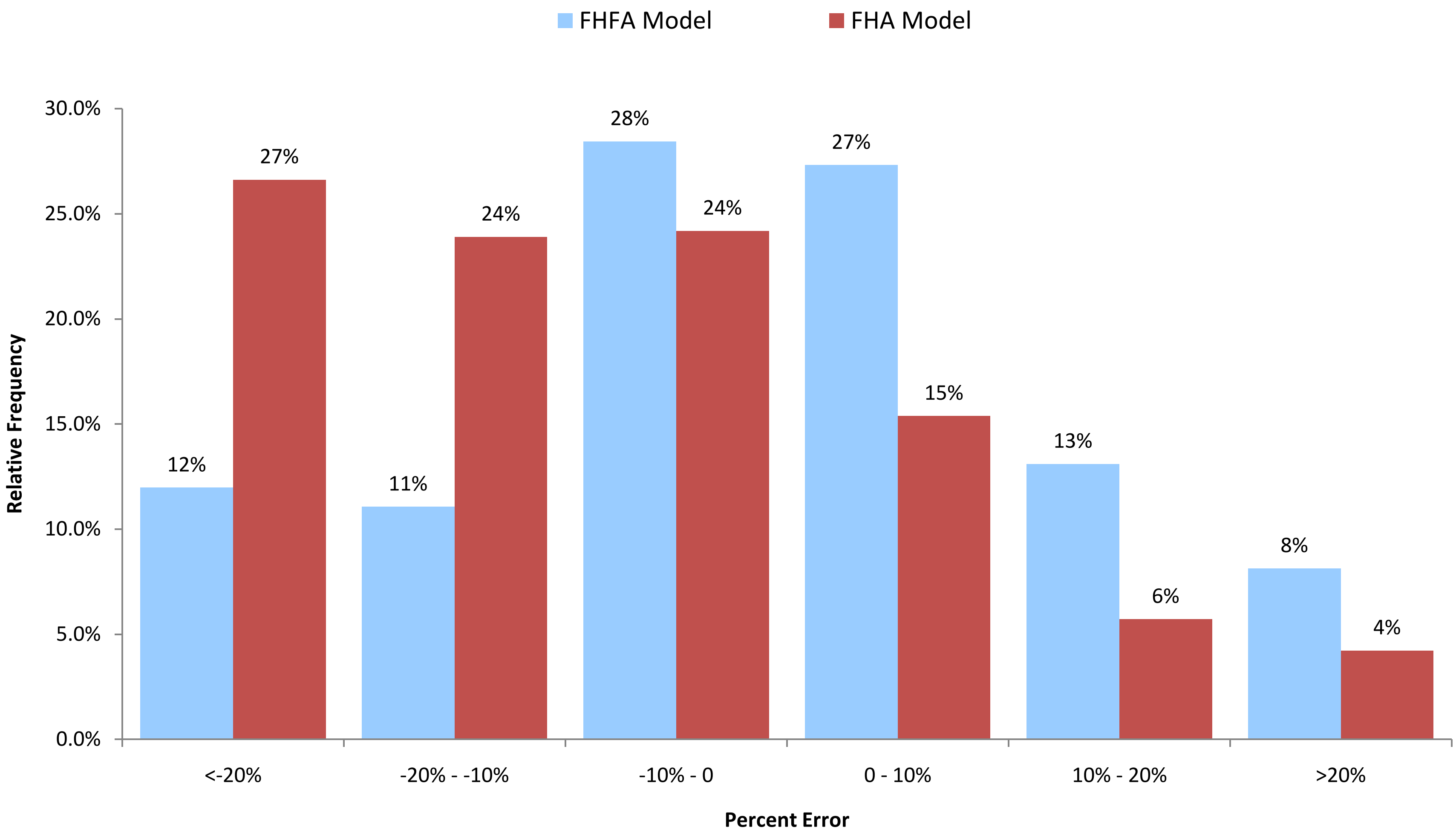
Figure 2: Relationship Between Relative Error and Price



Note: Relative error calculated as (Actual Price-Estimated Price)/Actual Price.

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, and First American CoreLogic ("LoanPerformance Data").

Figure 3: Relative Frequency of Prediction Error
County-By-County



Note: Relative error calculated as (Actual Price-Estimated Price)/Actual Price.

Sources: Datasets from DataQuick Information Systems, the Enterprises, FHA, RadarLogic, and First American CoreLogic ("LoanPerformance Data").