Highlights

Distress-Free House Price Indexes

Background

The Highlights article in the 2012Q1 HPI release noted that FHFA was evaluating various options for producing "distress-free" house price indexes. These indexes would remove the effect of short sales and real estate owned (REO) transactions (bank sales of foreclosed property) from the HPI. The article indicated that, in some situations, distress-free measures might be less noisy than the traditional HPI and might provide more relevant measures of changes in house prices.

A significant challenge in forming such indexes is finding a way of identifying distressed-sales in the transactions data. The transactions databases at FHFA's disposal—notably the Fannie Mae and Freddie Mac (the Enterprises) mortgage databases used for forming the traditional HPI—do not identify the name of the seller in real estate property transactions. It is thus difficult to know whether a given transaction represents an REO sale. Similarly, short sales cannot be identified consistently because the mortgage delinquency status and financial condition of the property seller is not always known.

FHFA has information about the status of the seller in select circumstances. For example, where the seller of a property financed the home with an Enterprise-guaranteed mortgage, Enterprise mortgage performance data can be used to identify cases where the sellers were late on their payments (and thus prone to short selling). Through a data sharing arrangement, FHFA also has mortgage performance data on certain FHA-endorsed mortgages. As such, mortgage distress can also be flagged in cases where the seller had an FHA-endorsed loan. Although the FHA and Enterprise mortgage performance data are helpful, unfortunately a sizeable proportion of the mortgage distress—and thus short sales—involve homes that were financed with other types of loans. Accordingly, FHFA's does not "see" many short sales with current data.

Last quarter's Highlights article focused on a method for identifying distressed sales that relied on information found in a new electronic appraisal database. While that approach was promising for identifying short sales and REO sales in the future, historical data limitations of the database meant that it could identify few distressed sales prior to late 2011.

Another approach—one that can identify historical distress sales extending back many more years—is used in this article. The methodology relies on several databases, including a licensed dataset of foreclosure-related filings recorded at county recorder offices. These data are used to identify distressed sales in twelve metropolitan areas in the United States. The twelve metropolitan areas include the ten cities with the largest

peak-to-current price declines (as measured with FHFA's purchase-only metropolitan area indexes) plus the Atlanta-Sandy Springs-Marietta, GA Metropolitan Statistical Area and the San Francisco-San Mateo-Redwood City Metropolitan Division.

The identified distress sales are removed from the HPI estimation data sample in these areas and, when estimated on this new dataset, FHFA's standard repeat-transactions methodology is used to produce distress-free house price measures. To study the impact of short sales and REO transactions on price measures, distress-free versions of FHFA's "purchase-only" indexes—metrics constructed using sales prices from Enterprise-guaranteed purchase-money mortgages—are then compared to FHFA's standard purchase-only measures.

Methodology

To identify short sales and REO sales in this analysis, three different databases are being used. The aforementioned mortgage performance data from the Enterprises and FHA comprise the first database. As indicated earlier, these data can be used to directly identify mortgage distress and REO sales where the sellers had Enterprise or FHA loans.

The second database includes information on foreclosure deeds recorded at county recorder offices. FHFA has licensed county deed recordations from DataQuick Information Systems for many counties throughout the country and, because the foreclosure process often culminates with certain types of deeds being recorded, deed information can be used to flag REO sales. For example, in California and Michigan, "Trustee Deeds Upon Sale" and "Sherriff's Deeds" are, respectively, used to convey property ownership after foreclosure auctions. Because, more often than not, banks take possession after foreclosure auctions, property transactions that occur after such deeds are recorded are likely to be bank sales.

The third database that can be used is a dataset FHFA recently licensed from CoreLogic. The dataset includes specific types of earlier-stage foreclosure filings that have been recorded at county recorder offices. In many jurisdictions across the country—including many counties in the twelve metropolitan areas analyzed here—certain types for formal notifications must be filed at the county recorder offices before the final phases of foreclosure can be completed. In California, for example, the first step in the foreclosure process involves the filing of a "Notice of Default" (NOD) which reports that a borrower is delinquent in his payments. In "judicial foreclosure" states where the foreclosure process is done through courts, *Lis Pendens* (LP) notices are filed to provide public notice of the existence of the foreclosure-related lawsuit.

This third dataset is particularly valuable because it can be used to identify short sales where sellers were in financial distress but did not have either an FHA or Enterprise mortgage. Because the notification-type filings (i.e., NODs and LPs) are registered for properties with all mortgage types—the dataset can be used to infer mortgage distress in cases where the seller had other types of financing (i.e., cases where FHFA's other datasets do not provide sufficient information).

In 2009, FHFA published an analysis that used NOD data to construct distress-free house price indexes in California.[1] The estimation strategy in that paper was simple and aligned with the approach used here: all property sales that occurred less than a year after an NOD filing were assumed to be short sales or REO transactions. After such observations were removed from the data sample, the index model was then re-estimated, thus forming a distress-free index.

Distress-free indexes are formed here using not only the earlier-stage foreclosure data, but also the other two databases identified above. The basic approach remains the same—all transactions that occur after some indication of distress are removed—but the new data sources are now used. Also, while the 2009 report focused on the state of California, the distress-free indexes reported here are for metropolitan areas in several different states.

To construct distress-free versions of FHFA's purchase-only indexes, transactions data are removed from the estimation data sample using a twelve-month rule. For a given transaction, if any of the three data sources report some type of distress associated with purchased property in the twelve months prior, then the transaction is flagged as "distressed" and removed. The specific indications of distress used include:

1. Enterprise or FHA mortgage delinquency was evident for the seller. The mortgage on the property is noted as having been "delinquent" if, at any point in the twelve months, the borrowers (i.e., the ultimate sellers) were two months or more delinquent on his payments.
2. Any one of the following was filed for the property in the preceding twelve months: a Trustee Deed Upon Sale, a Foreclosure Deed, a Sherriff's Deed, or Certificate of Final Judgment. As noted, these filings are associated with the conveyance of property ownership to banks in the final stage of foreclosure. Observed transactions that occur after such filings will tend to be REO sales.
3. Any of the following "notices" were filed: Notice of Default, Notice of Trustee Sale, or *Lis Pendens*.

[1] See "The Impact of Distressed Sales on Repeat-Transactions House Price Indexes," FHFA Research Paper published May 27, 2009 (http://www.fhfa.gov/webfiles/2916/researchpaper_distress%5b1%5d.pdf).

Foreclosure practices—and the precise deeds and other recordations that reflect short sales and foreclosure completions—vary across states. Indeed, they can vary across different counties within the same state. Given this heterogeneity, the myriad types of filings that are sometimes evident even within the same jurisdiction, and the fact that FHFA has only just begun using the early-stage foreclosure data from CoreLogic, it should be noted that the precise decision rules identified above are subject to modification. In releasing distress-free indexes in the future, FHFA may refine these rules.

Estimates

Using the 2012Q2 HPI data submissions from the Enterprises in conjunction with the data sources and methodology discussed above, Table 1 reports the share of distressed transactions in the HPI data sample over the last six quarters. These shares reflect the proportion of Enterprise purchase-money mortgages that financed homes that were sold by a seller who was in financial distress or was a bank (or other third party buyer of a foreclosed home).

Market observers have suggested that the share of distressed transactions sinks in the springtime as sales volumes for the "nondistressed" marketplace hit their seasonal highs. For the twelve cities analyzed here, Table 1 shows that the distressed sale share in the Enterprise sample generally fell between the first and second quarters in the last two years. This result was not evident for each of the twelve cities—some cities saw the increased shares. For the twelve cities in aggregate, however, it was true. Between the first and second quarters of 2012, for instance, the aggregate share of distressed sales in the twelve cities fell from 38.3 percent to 29.1 percent.

In general, the share of distressed sales in the Enterprise data sample (and undoubtedly in marketplace as a whole) is substantial and thus it is not surprising that distress-free house price indexes, which are made available for download here, can differ significantly from the standard "full-sample" indexes. Table 2 provides summary statistics that compare, by city, these new distress-free indexes against the standard purchase-only indexes. The table compares price changes for the latest quarter, the last four quarters, and the last five years.

Consistent with the fact that the share of the distressed sales fell somewhat between the first and second quarters, the quarterly price changes for the distress-free indexes generally show smaller price increases than the standard purchase-only indexes. In effect, some of the measured price increases over the quarter appear to be the result of the decline in the distressed sale share. For the Atlanta metropolitan area, for example, a 10 percentage point decrease in the share of the distress appears to explain some of the 5.2

percent price increase. The distress-free index shows only a 3.0 percent price rise over the quarter.

While the gap between the respective measurements of four-quarter price change estimates can also be significant, Table 2 reveals a relatively modest difference in the five-year price change estimates. It seems that quarter-to-quarter variations in the share of distressed sales drive some short-term price movements, but over the longer term, the respective measures do not diverge dramatically. Except for Atlanta, where the gap between the measured price declines is a significant 12.6 percentage points, the divergence is between 3 and 8 percentage points.

When comparing price changes reflected in the distress-free indexes against those in the full-sample metric, it should be noted that differences in the geographic mix of the respective samples may also explain part of gap between the measures. When distressed sales are removed from the data sample to form the distress-free metrics, there is an impact of the geographic representativeness of the data sample. The representation of specific sub-areas *within* each of the cities is altered when distressed sales are removed. Neighborhoods that saw the greatest numbers of distressed sales—neighborhoods which tended to have the largest price declines—will have smaller representation in the "distress-free" data sample. This means that, all else equal, the distress-free index will tend to show more modest price declines not only because they remove the direct effects of distressed sales, but also because they give less weight to price trends in neighborhoods with the greatest price declines.

Commentary

The distress-free indexes published in this Highlights article ought to be considered developmental in nature. FHFA intends to continue publishing such metrics with future HPI releases, but as noted, methodological refinements may be made. These enhancements likely will be focused on improving the accuracy with which distressed sales are identified.

FHFA would like to ensure that the indexes are constructed in a way that maximizes their usefulness for research and modeling purposes. Accordingly, FHFA welcomes public input. Comments, questions, and suggestions should be addressed to hpihelpdesk@fhfa.gov.

Table 1: Share of Distressed-Sales in HPI Data Sample
Fraction of Purchase-Money Mortgages Occurring after Distress Indication

Metropolitan Area	2011Q1	2011Q2	2011Q3	2011Q4	2012Q1	2012Q2
Atlanta-Sandy Springs-Marietta, GA	33.2%	34.1%	34.2%	37.8%	38.7%	28.3%
Chicago-Joliet-Naperville, IL (MSAD)	26.0%	21.4%	17.0%	21.7%	24.9%	17.1%
Los Angeles-Long Beach-Glendale, CA (MSAD)	39.5%	36.7%	34.9%	36.5%	38.8%	32.3%
Miami-Miami Beach-Kendall, FL (MSAD)	38.4%	30.8%	24.3%	22.2%	28.8%	16.6%
Oakland-Fremont-Hayward, CA (MSAD)	45.6%	39.8%	37.8%	41.0%	45.7%	33.2%
Phoenix-Mesa-Glendale, AZ	60.5%	62.0%	60.6%	55.6%	49.2%	40.4%
Riverside-San Bernardino-Ontario, CA	62.0%	61.0%	60.3%	57.7%	59.3%	51.8%
San Diego-Carlsbad-San Marcos, CA	40.0%	35.8%	35.2%	38.2%	36.8%	31.5%
San Francisco-San Mateo-Redwood City, CA (MSAD)	26.9%	21.7%	22.6%	25.3%	28.5%	18.9%
Santa Ana-Anaheim-Irvine, CA (MSAD)	34.9%	32.2%	31.9%	35.2%	36.7%	29.2%
Tampa-St. Petersburg-Clearwater, FL	35.0%	30.0%	25.9%	20.8%	21.6%	18.6%
Warren-Troy-Farmington Hills, MI (MSAD)	31.2%	26.3%	19.2%	22.7%	26.2%	18.7%
All 12 Metropolitan Areas	41.6%	37.7%	35.1%	36.7%	38.3%	29.1%

Sources: Fannie Mae and Freddie Mac mortgage data, including mortgage performance records; FHA mortgage performance data; county recorder data from DataQuick Information Systems; Notice of Default, *Lis Pendens* and other foreclosure-related filings data licensed from CoreLogic

Table 2: Price Changes Measured with Purchase-Only HPI
Full Sample vs. Distress-Free Indexes
(All Estimates are Seasonally Adjusted)

Metropolitan Area	Quarterly Change (2012Q1-2012Q2)		Four-Quarter Change (2011Q2-2012Q2)		Five-Year Change (2007Q2-2012Q2)	
	Full Sample	Distress-Free	Full Sample	Distress-Free	Full Sample	Distress-Free
Atlanta-Sandy Springs-Marietta, GA	5.2%	3.0%	6.3%	1.1%	-26.0%	-13.4%
Chicago-Joliet-Naperville, IL (MSAD)	1.5%	1.0%	-0.7%	-0.8%	-29.7%	-22.1%
Los Angeles-Long Beach-Glendale, CA (MSAD)	5.3%	3.1%	2.5%	0.5%	-35.4%	-27.8%
Miami-Miami Beach-Kendall, FL (MSAD)	8.3%	8.9%	11.0%	10.1%	-41.6%	-37.4%
Oakland-Fremont-Hayward, CA (MSAD)	3.1%	2.0%	5.1%	1.1%	-38.4%	-30.3%
Phoenix-Mesa-Glendale, AZ	6.9%	4.6%	13.9%	3.6%	-44.7%	-40.5%
Riverside-San Bernardino-Ontario, CA	2.3%	-1.6%	4.2%	-1.0%	-49.5%	-47.1%
San Diego-Carlsbad-San Marcos, CA	3.7%	3.1%	5.6%	5.5%	-27.3%	-21.3%
San Francisco-San Mateo-Redwood City, CA (MSAD)	3.5%	4.2%	4.5%	3.4%	-16.4%	-12.3%
Santa Ana-Anaheim-Irvine, CA (MSAD)	-1.2%	-2.0%	-2.6%	-3.8%	-30.7%	-26.7%
Tampa-St. Petersburg-Clearwater, FL	3.3%	4.4%	7.3%	7.3%	-36.8%	-29.6%
Warren-Troy-Farmington Hills, MI (MSAD)	4.3%	1.6%	10.0%	3.8%	-25.7%	-22.5%

Sources: Fannie Mae and Freddie Mac mortgage data, including mortgage performance records; FHA mortgage performance data; county recorder data from DataQuick Information Systems; Notice of Default, *Lis Pendens* and other foreclosure-related filings data licensed from CoreLogic

Technical Note

A Change in the "Sales Price" Data Used to Estimate the FHFA HPI

In this release, a small but notable change has been made to how the HPI is calculated. As in prior periods, the indexes continue to be constructed using house price data from loans bought or guaranteed by Fannie Mae and Freddie Mac (the Enterprises). The change involves the price measure that is used for purchase-money mortgages.[1]

Prior to this period, the "sales price" used in connection with a purchase-money mortgage was a calculated value based on a mortgage's loan-to-value (LTV) ratio and the loan amount The loan amount was divided by the LTV ratio and the result was used as the measure of the actual selling price.

The imputation approach stemmed from the fact that, when the HPI was first constructed in the mid-1990s, sales price information was generally not available for most historical mortgages in the Enterprise data systems. Until recently, the Enterprises' HPI data submissions to FHFA did not include a sales price field because of the prevalence of missing values in early periods.

Although very close to actual sales prices in the vast majority of cases, the imputed value was an imperfect measure for two reasons. The primary source of potential divergence between its value and the actual sales price was the fact that the LTV ratio was not simply the ratio of the loan amount to sales price (if it were, the estimated value would always be equal to the sales price). The LTV ratio—a value submitted to the Enterprises by loan originators—represents the loan amount divided by the *lesser of* the sales price and the appraised value. Where appraised home values were the same as or above the sales price—which was the case for the most mortgages until very recently—the use of the "lesser of" rule had no effect: the imputed value was the actual selling price in those cases. Where the appraised value was below the selling price, however, the use of the "lesser of" rule meant that the imputed value would be below the property sales price.

The second, much smaller issue was that the LTV ratio in the Enterprises' data systems is a *rounded* value, as opposed to the precise ratio of the loan amount to the home value. The rounding meant that, even if the sales price was below the appraised value (and thus the "lesser of" rule was not a problem), the imputed sales price was

[1] The home values that are used in connection with refinance mortgages remain the appraised values. Such values continue to be used only in the construction of FHFA's "all-transactions" indexes.

sometimes different than the actual sales price. Although producing no systemic bias, this generated noise in index estimation.

To avoid this noise and the problems caused by the "lesser of" rule, with this release, FHFA has begun using the property sales price as reported in the Enterprises' data submissions. Beginning in the summer of 2011, the Enterprises began reporting this field to FHFA in its monthly data submissions and, since then, FHFA has evaluated the benefits and disadvantages associated with using it rather than the imputed value. On balance, FHFA believes that using the sales price variable instead of the imputed value is warranted.

Unfortunately, unlike imputed home values—which can be formed for every mortgage transaction—sales prices are unavailable for some mortgages. The share of missing values is particularly pronounced for mortgages originated in the early 1990s and before, when very few loans have recorded selling prices. To ensure that these data points are not dropped from the estimation sample, in implementing the model change, FHFA has decided to use the sales price field only where: (1) the mortgage was originated in 1995 or later AND (2) the reported sales price is not missing.

The changeover to the sales price field has a very modest impact on HPI estimates. For states and the U.S. as a whole, Table 1 shows the effect on quarterly price estimates reflected in the purchase-only index. The table also shows the impact on the index *levels* for those geographic areas.

Both metrics reveal modest effects. The price change estimates determined under the new price measure are neither systematically above nor below estimates that are produced under the old measure. The average absolute difference in the seasonally adjusted quarterly change for the 50 states and Washington, D.C. is near zero (0.1 percentage points).

The impact on the index *values* is also modest. The "old" and "new" index values differ by less than one index point (less than 0.5 percent of index levels) in the majority of cases. The largest difference is for Hawaii, where the difference is a still-modest 1.8 index points.

Technical Note

Table 1: Impact of Using "Sales Price" Instead of Imputed Sales Price
(Purchase-Only, Seasonally Adjusted HPI)

	Quarterly Price Change (2012Q1-2012Q2)			Index Level (2012Q2)		
	New Approach: Use of Sales Price Field	Old Approach: Imputed Sales Price	Difference	New Approach: Use of Sales Price Field	Old Approach: Imputed Sales Price	Difference
United States	1.8%	1.8%	0.0%	185.5	184.8	0.6
Alabama	1.2%	1.2%	0.0%	179.2	178.5	0.7
Alaska	4.6%	4.6%	-0.1%	227.3	226.7	0.6
Arizona	6.0%	6.0%	0.0%	181.9	180.8	1.1
Arkansas	1.9%	1.8%	0.1%	185.9	185.3	0.6
California	3.5%	3.5%	0.1%	158.9	158.1	0.8
Colorado	4.0%	4.0%	0.0%	270.4	269.5	0.9
Connecticut	-1.4%	-1.2%	-0.2%	161.9	161.6	0.4
Delaware	-0.6%	-0.2%	-0.4%	169.4	169.4	0.0
District of Columbia	-0.1%	-0.8%	0.7%	358.0	356.6	1.4
Florida	3.3%	3.4%	-0.1%	180.4	179.6	0.8
Georgia	2.6%	2.6%	0.1%	152.4	151.8	0.6
Hawaii	2.9%	2.7%	0.2%	183.1	181.3	1.8
Idaho	3.9%	3.7%	0.2%	195.8	195.2	0.5
Illinois	1.6%	1.6%	0.0%	171.8	171.2	0.6
Indiana	0.7%	0.7%	-0.1%	159.6	159.3	0.2
Iowa	-1.1%	-0.5%	-0.6%	197.6	197.3	0.3
Kansas	0.3%	-0.1%	0.4%	190.6	190.0	0.6
Kentucky	1.8%	1.6%	0.2%	190.7	189.6	1.1
Louisiana	1.7%	1.8%	-0.1%	228.6	228.0	0.6

Technical Note

Table 1: Impact of Using "Sales Price" Instead of Imputed Sales Price
(Purchase-Only, Seasonally Adjusted HPI)

	Quarterly Price Change (2012Q1-2012Q2)			**Index Level** (2012Q2)		
	New Approach: Use of Sales Price Field	Old Approach: Imputed Sales Price	Difference	New Approach: Use of Sales Price Field	Old Approach: Imputed Sales Price	Difference
Maine	0.4%	0.3%	0.1%	203.2	202.7	0.5
Maryland	4.7%	4.9%	-0.2%	210.9	210.4	0.4
Massachusetts	0.5%	0.5%	0.0%	214.4	213.9	0.5
Michigan	3.5%	3.6%	-0.1%	149.2	148.9	0.4
Minnesota	1.8%	1.8%	0.0%	204.8	204.4	0.5
Mississippi	-1.6%	-1.6%	0.0%	174.6	173.6	1.0
Missouri	0.8%	0.8%	0.0%	182.7	182.2	0.5
Montana	0.5%	0.4%	0.1%	292.6	290.9	1.7
Nebraska	0.8%	0.8%	0.0%	196.2	195.7	0.4
Nevada	4.8%	4.8%	0.0%	116.8	116.1	0.7
New Hampshire	1.4%	1.3%	0.0%	194.3	193.5	0.7
New Jersey	1.6%	1.5%	0.2%	211.6	210.7	0.9
New Mexico	3.2%	2.9%	0.3%	210.9	209.5	1.3
New York	0.1%	0.0%	0.1%	204.5	203.8	0.7
North Carolina	0.2%	0.2%	0.0%	178.4	178.1	0.2
North Dakota	1.8%	2.1%	-0.3%	242.7	242.5	0.2
Ohio	0.9%	1.1%	-0.1%	153.6	153.2	0.4
Oklahoma	0.3%	0.4%	-0.1%	194.2	193.5	0.6
Oregon	2.0%	2.0%	0.0%	253.4	252.8	0.6
Pennsylvania	0.6%	0.7%	-0.1%	186.7	186.2	0.5
Rhode Island	-0.7%	-0.7%	0.0%	179.2	178.3	0.9

Technical Note

Table 1: Impact of Using "Sales Price" Instead of Imputed Sales Price
(Purchase-Only, Seasonally Adjusted HPI)

	Quarterly Price Change (2012Q1-2012Q2)			Index Level (2012Q2)		
	New Approach: Use of Sales Price Field	Old Approach: Imputed Sales Price	Difference	New Approach: Use of Sales Price Field	Old Approach: Imputed Sales Price	Difference
South Carolina	1.9%	1.8%	0.1%	178.9	178.2	0.7
South Dakota	2.6%	2.7%	-0.1%	229.8	229.6	0.2
Tennessee	1.8%	1.9%	-0.1%	185.4	184.9	0.5
Texas	1.3%	1.2%	0.1%	195.7	195.0	0.7
Utah	2.5%	2.5%	0.0%	254.2	253.2	1.0
Vermont	-0.9%	-1.3%	0.4%	209.0	207.9	1.1
Virginia	0.8%	0.8%	0.1%	211.5	210.8	0.7
Washington	3.6%	3.7%	-0.1%	213.2	212.7	0.5
West Virginia	-3.7%	-3.3%	-0.4%	189.3	188.9	0.3
Wisconsin	0.8%	0.8%	0.0%	203.1	202.6	0.5
Wyoming	1.7%	1.4%	0.2%	289.7	290.0	(0.4)