WORKING PAPERS

RESEARCH DEPARTMENT

WORKING PAPER NO. 14-38
UNDERSTANDING HOUSE PRICE INDEX REVISIONS

Ronel Elul
Federal Reserve Bank of Philadelphia

Joseph M. Silverstein
University of Pennsylvania

Tom Stark
Federal Reserve Bank of Philadelphia

December 2014

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA

Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Understanding House Price Index Revisions

Ronel Elul*

Joseph M. Silverstein

Tom Stark

November 1, 2014

Abstract

Residential house price indexes (HPI) are used for a large variety of macroeconomic and microeconomic research and policy purposes, as well as for automated valuation models. As is well known, these indexes are subject to substantial revisions in the months following the initial release, both because transaction data can be slow to come in, and as a consequence of the repeat sales methodology, which interpolates the effect of sales over the entire period since the house last changed hands.

We study the properties of the revisions to the CoreLogic House Price Index. This index is used both by researchers and in the Financial Accounts of the United States to compute the value of residential real estate. We show that the magnitude of revisions to this index can be significant: At the national level, the ratio of standard deviation of monthly revisions to the growth rate of the index, relative to the standard deviation of the growth rate in the index, is 29%, which is comparable to the relative ratio for other macroeconomic series. The revisions are also economically significant and impact measures used by policymakers: Revisions over the first 12 releases of the index reduce estimates of the fraction of borrowers nationwide with negative equity by 4.3%, corresponding to 423,000 households. Lastly, we find that revisions are ex-ante predictable: Both past revisions and past house price appreciation are negatively correlated with future revisions.

Keywords: house price index, data revisions, real-time data

JEL Codes: R21, R31

* Elul and Stark are both from the Federal Reserve Bank of Philadelphia. This research began while Joseph Silverstein was a research assistant at the Federal Reserve Bank of Philadelphia. We thank George Alessandria, Dean Croushore, Will Doerner, Nick Reynolds, and Joe Tracy for helpful discussions. We also thank Will Doerner for providing us with data on revisions to the FHFA index. Contact: Ronel Elul (ronel.elul@phil.frb.org, 215-574-3965).

Introduction

Residential house price indexes (HPI) are used for a large variety of macroeconomic and microeconomic research and policy purposes, as well as for automated valuation models. As is well known, these indexes are subject to substantial revisions in the months following the initial release, both because transaction data can be slow to come in, and as a consequence of the repeat sales methodology, which interpolates the effect of sales over the entire period since the house last changed hands.

We study the properties of the revisions to the CoreLogic House Price Index (HPI). This index is used widely, both by researchers and, since 2008, in the Flow of Funds to compute the value of residential real estate. We show that the revisions to this index can be significant: At the national level, the ratio of standard deviation of monthly revisions in the growth rate of index, relative to the standard deviation of the growth rate, is 29%. Moreover, these revisions can also be economically significant, and impact measures used by policymakers: We find that revisions over the first 12 releases of the HPI data reduced estimates of the fraction of borrowers nationwide with negative equity by 4.3%, on average, with substantially larger drops in some time periods.

Lastly, we demonstrate that revisions are predictable, in that past revisions and house price appreciation are negatively associated with future revisions. In addition, we find that these revisions are positively correlated with contemporaneous changes in house prices. This later finding suggests, in particular, that initial data on the impact of programs to stimulate housing markets may tend to underestimate their total effect.

Characteristics of Revisions

We use data on house price indexes and their revisions from CoreLogic, Inc., from March 2011 through February 2014.[1] This index uses the arithmetic repeat-sales methodology as in Shiller (1991). The index is available for various levels of aggregation and sales tiers: We consider, in particular, the single-family detached index at the national, state, and zip code levels. Monthly revisions of the index are also available for this time span. Similar methodology is used

[1] See Tracy et al. (2013) for further discussion of revisions to the CoreLogic index.

in constructing the Office of Federal Housing Enterprise Oversight (OFHEO) index; Deng and Quigley (2008) study the characteristics of this index.

Figure 1 gives the initial release of the CoreLogic index at the national level, along with the 12th release of the index and the latest release (201402).

In Figure 2, we plot average values of the monthly marginal revisions to the level of the CoreLogic index (i.e., the revision relative to the previous release), from the first through the 15th revision, at the national, state, and zip code levels. Not surprisingly, revisions are smaller at higher levels of aggregation. The mean first revision for the U.S. is 0.19%, and for the states, the average is -0.07%. Later revisions decline in magnitude; for the national index, the third revision is approximately one-third the size of the first revision.

To gauge the relative size of these revisions, in Table 1, we present the mean and standard deviation of the growth rate of the HPI and its first revisions. At the national level, the ratio of the SD of first revisions to the growth rate, relative to the SD of the one-month growth rate is 0.29.[2] This is comparable to the ratio for revisions to other macroeconomic series. For example, the relative standard deviation for the first monthly revision to nonfarm payrolls is approximately 0.3, and the ratio for the first revision to the growth rate of real GDP is 0.25.[3] At the state level the relative ratios for revisions to the HPI growth rate are larger, averaging 0.63.

To give a better idea of the magnitude of these revisions, Figure 3 shows the average absolute values of the monthly marginal revisions at the national, state, and zip code levels. At lower levels of aggregation, we find substantial variation in the magnitude of revisions: For example, the average absolute value of first revisions in Colorado is 0.29%, while the average absolute value of first revisions in West Virginia is 4.15%. In Figure 4, we break up the states into five groups, based on the absolute value of the initial marginal revision, and then plot the magnitude of subsequent revisions; we show that those states with larger initial revisions also tend to have larger subsequent ones.[4]

---

[2] By way of comparison, the relative ratio for the first quarterly revision to the Federal Housing Finance Agency (FHFA) index at the Census division level is 0.11.

[3] We compute the ratio of the standard deviation of the first revision to the growth rate of the HPI, normalized by the standard deviation of the growth rate of the latest release of the index. The data used to generate the comparable ratios for the change in nonfarm payrolls and GDP growth rate are from the Philadelphia Fed's Real-Time Data Set for Macroeconomists (http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/).

[4] Q1 (smallest 1st revision): AZ, CA, CO, IA, NV, OK, OR, PA, TN, TX, WA; Q2: AR, FL, KY, NJ, NY, NC, OH, UT, VA, WI; Q3: AK, GA, ID, IL, IN, LA, MD, MO, NE, NM; Q4: AL, CT, KS, MA, MI, MN, NH, RI, SC, SD; Q5: DC, DE, HI, ME, MS, MT, ND, VT, WV, WY. By examining this list of state groups, we can see that one contributor to this persistence in magnitude is the size of the states: smaller states tend to have larger revisions.

Although we later show that there is some negative autocorrelation in revisions, nevertheless revisions are persistent, as can be seen from Figure 5, which plots the cumulative revisions; these are monotonically increasing.

We also provide histograms of first revisions at the national and state level in Figure 6.

Some Notation

Let $P_{t,i}^{v(j)}$ denote the HPI at date *t*, for state *i*, as of release date (i.e., "vintage") *j*. Thus, the initial release of the data would be given by $P_{t,i}^{v(t)}$, the second release (i.e., first revision) by $P_{t,i}^{v(t+1)}$, and so on. And the first (percent) cumulative revision to the date-*t* index is $\frac{P_{t,i}^{v(t+1)}}{P_{t,i}^{v(t)}} - 1$ and, more generally, the m^{th} cumulative revision (i.e., relative to the initial release) is denoted by $\frac{P_{t,i}^{v(t+m)}}{P_{t,i}^{v(t)}} - 1$.[5] Similarly, the marginal revisions would be defined relative to the previous release. We will also consider the initial release of the one-month growth rate in the index: $\frac{P_{t+1,i}^{v(t+1)}}{P_{t,i}^{v(t+1)}} - 1$; revisions to growth rates will be defined analogously.

The Economic Significance of HPI Revisions

To date, little work has been done on the economic significance of HPI revisions; this paper is one of the first to examine this issue.[6] To gauge the significance of HPI revisions, we consider their impact on the estimated share of mortgages with negative equity as well as those with loan-to-value ratios (LTV) above 80%. These populations are of interest because they are associated with elevated mortgage default rates (see Elul et al, 2010). In addition, one might be interested in the size of the eligible population for mortgage refinancing programs such as HARP that target these high-LTV homeowners.

We work with a 3% random sample of active first mortgages from the LPS dataset. The LPS data set covers approximately two-thirds of all outstanding mortgages and reports the origination date, balance at origination, and the value of the property on the origination date (the

[5] Note that our convention here is to date vintages using the date of the last observation in the vintage.
[6] Some, such as Deng and Quigley (2008), do however consider the effect that revisions might have on HPI derivatives.

purchase price, or appraised value in the case of refinancings), as well as the zip code in which the property is located. For subsequent months, we obtain the updated current principal balance.[7]

For each of these mortgages, we estimate the current loan-to-value ratio in each month of our sample by taking the first mortgage principal balance at that time and dividing by an updated estimate of the house value. The latter is obtained by taking the house value as reported at the time of mortgage origination and applying the change in the CoreLogic HPI for the property zip code from the origination date up through to that time. We do this for the initial release of the HPI for the month in question, as well as for subsequent revisions to the index, and focus, in particular, on the cumulative change in the index through the 12th release of the HPI data.

Table 2 confirms that index revisions have a significant impact on the share of the population with negative equity. The share of mortgages with negative equity — based on the initial HPI release — average 21% in our sample. In addition, for the period of March 2011–February 2014, the CoreLogic index[8] was revised up by 1.11%, on average, in the year following the initial release. These index revisions are then associated with a downward revision of 0.81 percentage points in the share of mortgages with negative equity, on average. Scaling our sample estimates to the entire U.S. mortgage market, this implies a drop in the number of households with negative equity by 423,000, on average, in each month, solely as a result of index revisions. Moreover, in some months (late 2011 and early 2012), the impact of revisions is substantially larger. This significant impact on the negative equity share occurs because these revisions were largest in those areas with the greatest share of households with negative equity, as shown in Figure 7. For instance, one-year house price revisions average 1.68% for state-date pairs in the top quartile of negative equity, as compared with -0.64% for those in the bottom quartile.

In Table 2, we also summarize the impact of revisions on the share of homeowners with LTVs at or above 80%; the results are similar to those reported for the negative equity share.

Predictability of Revisions

It is interesting to study whether the revisions are predictable. That is, can we use past releases of the index, as well as other information, such as past house price appreciation, to forecast future revisions? For the OFHEO index, similar work has been undertaken by Deng and

[7] For more information on the LPS data set, see Elul et al. (2010).
[8] This refers to the zip code level index, weighting by the share of active loans in the sample in that zip code.

Quigley (2008), who find some evidence of negative correlation between past appreciation and future revisions.

We first consider whether past revisions predict future ones. We run the following model at the state level:

$$\frac{P_{t,i}^{v(t+m)}}{P_{t,i}^{v(t)}} - 1 = \alpha + \sum_{j=1}^{4} \beta_j \cdot \left[\frac{P_{t-j,i}^{v(t-j+1)}}{P_{t-j,i}^{v(t-j)}} - 1 \right].$$

The results for first revisions (*m*=1) can be found in Table 3. We see first (model I) that revisions appear to have mean zero over this period (more generally, the constant term insignificant for almost all of our models). Models II and III add the lagged house price revisions; the first lag is always negative and significant. These regressions provide evidence that revisions tend to exhibit negative serial correlation, with positive revisions predicting negative ones in the future (and thus subsequent revisions partially offset each other).

We also use past house price appreciation as predictors:

$$\frac{P_{t,i}^{v(t+m)}}{P_{t,i}^{v(t)}} - 1 = \alpha + \sum_{j=1}^{4} \beta_j \cdot \left[\frac{P_{t-j+1,i}^{v(t)}}{P_{t-j,i}^{v(t)}} - 1 \right].$$

For *m*=1, the results are again in Table 3. Similar to lagged revisions, we find (models IV and V) that lagged house price appreciation is negatively associated with future revisions. This result is also consistent with Deng and Quigley (2008).

In models VI and VII, we include both lagged revisions and appreciation, and the results remain as above. We also see from the change in R^2 across the models that lagged revisions and house price appreciation make roughly similar contributions to the explanatory power of the model.

Models VIII–XV demonstrate that these results are robust to the inclusion of state and time fixed effects. We also see that the explanatory power of these fixed effects is of roughly the same order of magnitude as the combination of lagged revisions and appreciation in the earlier models.

Table 4 presents the estimation results from these models for longer-term revisions (*m*=3 and *m*=11).

We also study the relationship between revisions and contemporaneous house price appreciation; the results are in Table 5. We find (weak) evidence of a positive correlation between these two covariates. Observe that this is consistent with our results on lagged revisions and appreciation in models II and IV above, as both were negatively associated with future revisions. Although not statistically significant in our sample, it may be economically important, as it suggests that the impact of policies designed to stimulate the housing market may not be captured when they first begin to take effect. In this regard, our finding is reminiscent of Dynan and Elmendorf (2001).

Conclusion

We have characterized the properties of revisions to the CoreLogic HPI. We have shown that these revisions can be economically significant, and, in particular, that they impact estimates of key measures, such as the share of households with negative equity. We also demonstrate that revisions are predictable, with past revisions and house price appreciation being negatively associated with future revisions.

Areas for future study include the determinants of the magnitude of revisions across states, as well as further investigation of the predictability of these revisions.

Figures and Tables

Figure 1
HPI and Its Revisions: National Index

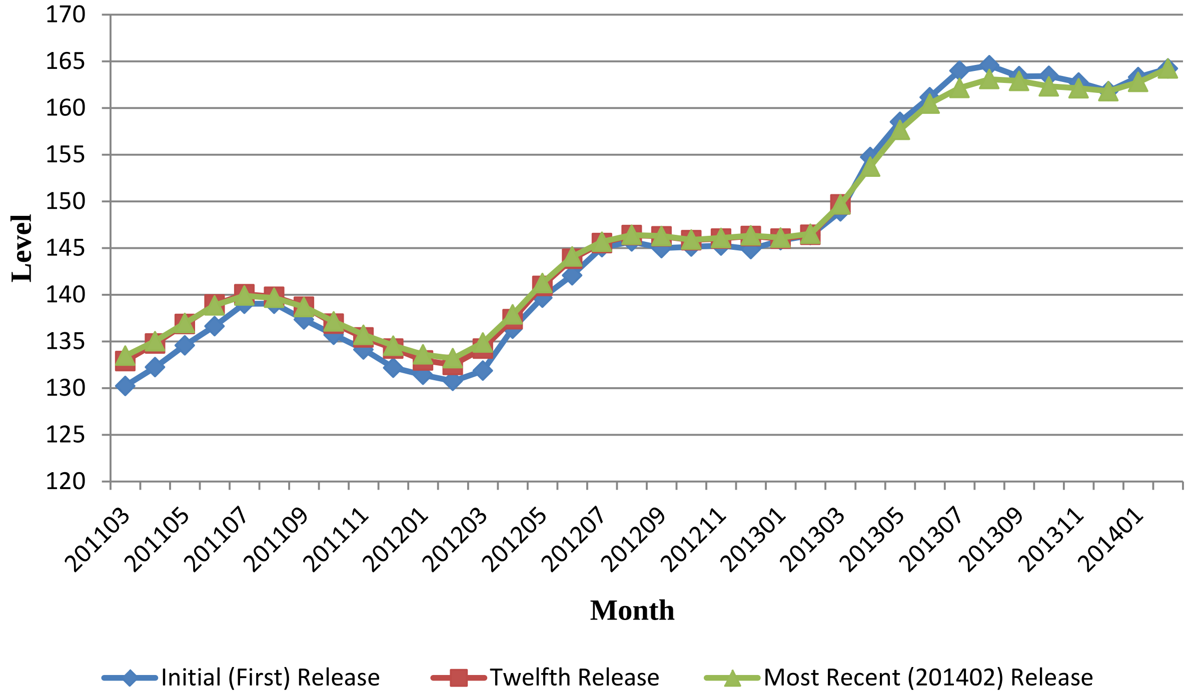


Figure 2
Average Monthly Revisions (%)

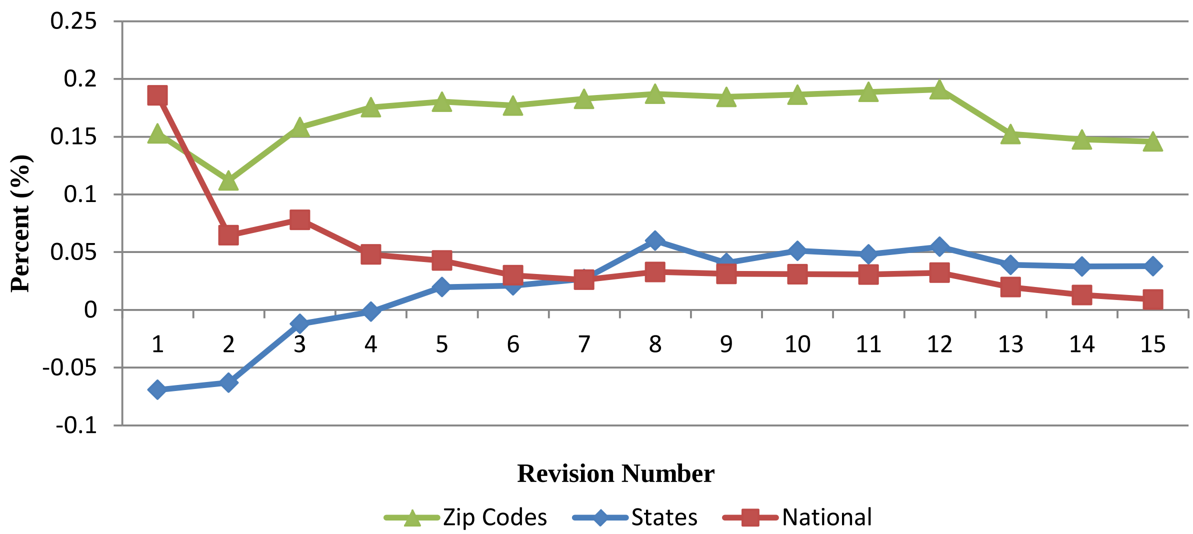
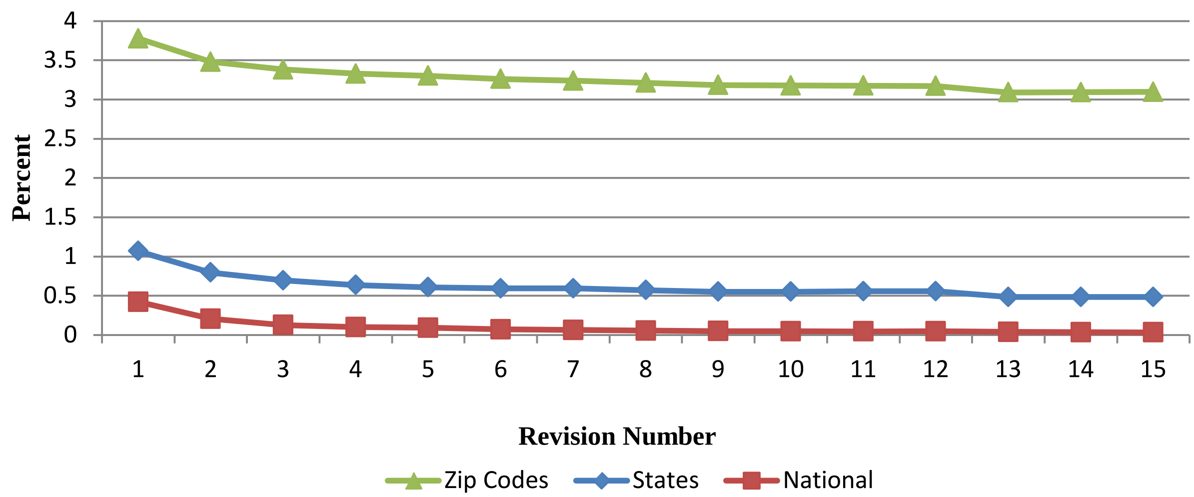


Table 1: HPI Growth Rate and Revisions — Summary Statistics

	National		State (Avg.)	
	Hpi Apprec.	1st Revision	Hpi Apprec.	1st Revision
Mean	0.59%	0.10%	0.35%	-0.04%
SD	1.10%	0.31%	1.14%	0.71%

Figure 3
Average Absolute Value of Monthly Revisions (%)



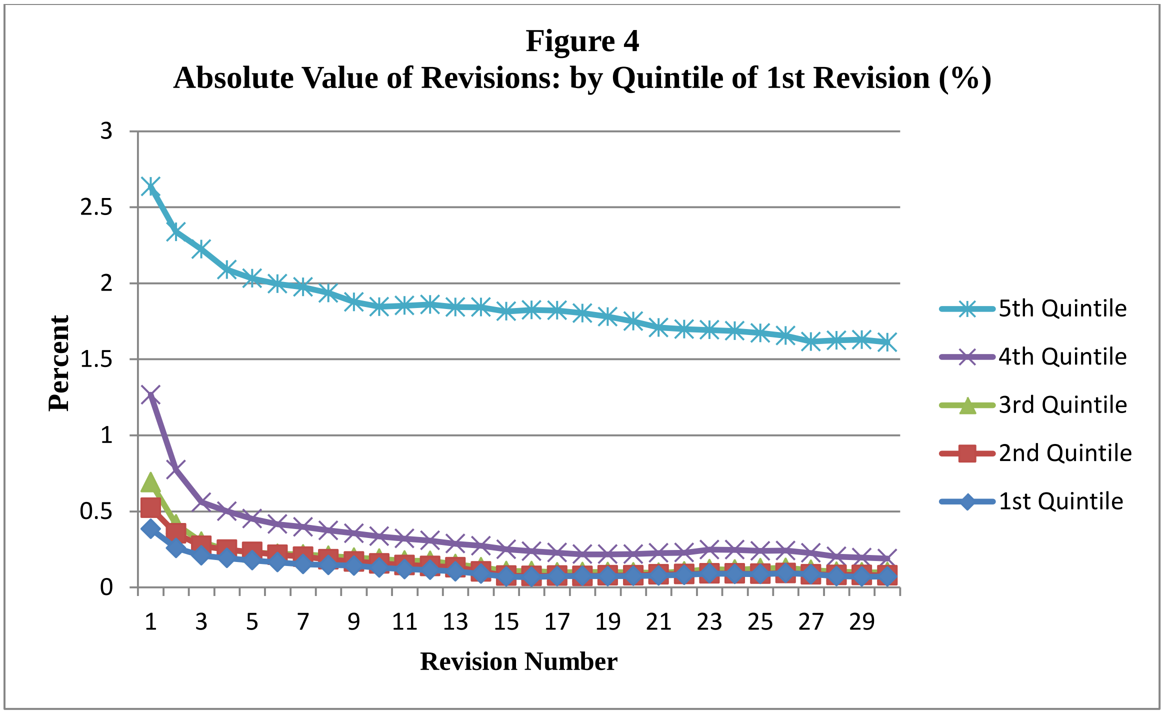


Figure 4
Absolute Value of Revisions: by Quintile of 1st Revision (%)
3
2.5
2
1.5
1
0.5
0
Percent
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
Revision Number
5th Quintile
4th Quintile
3rd Quintile
2nd Quintile
1st Quintile

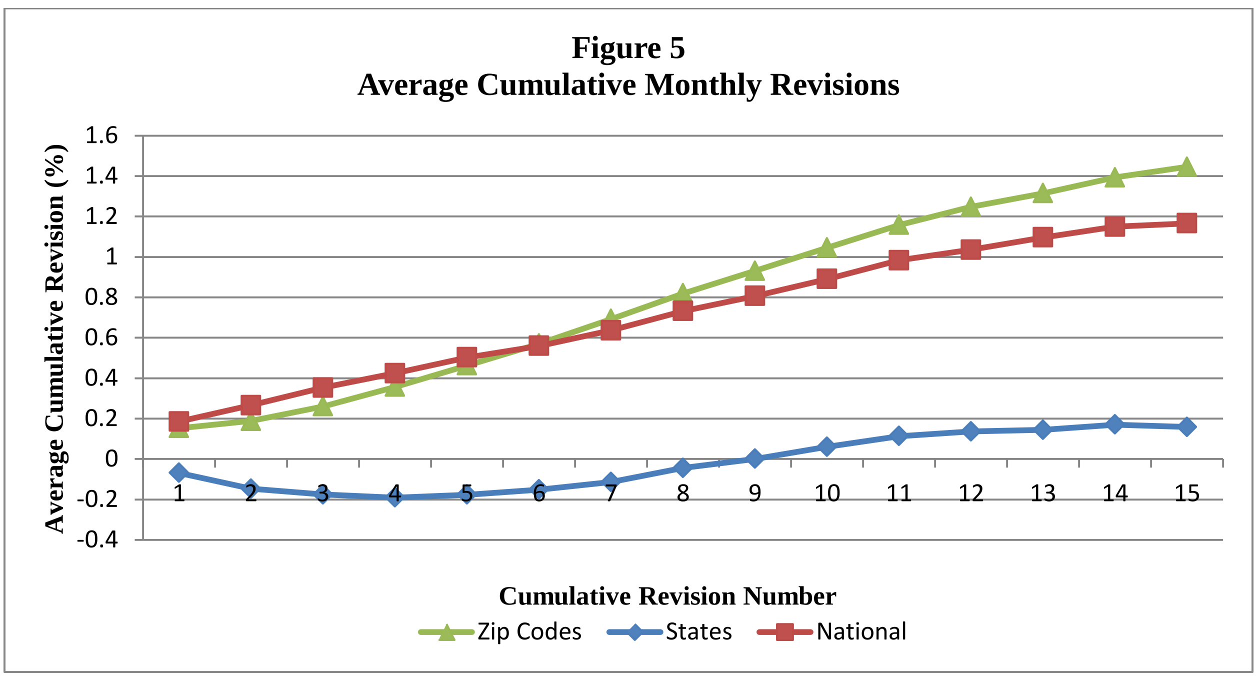
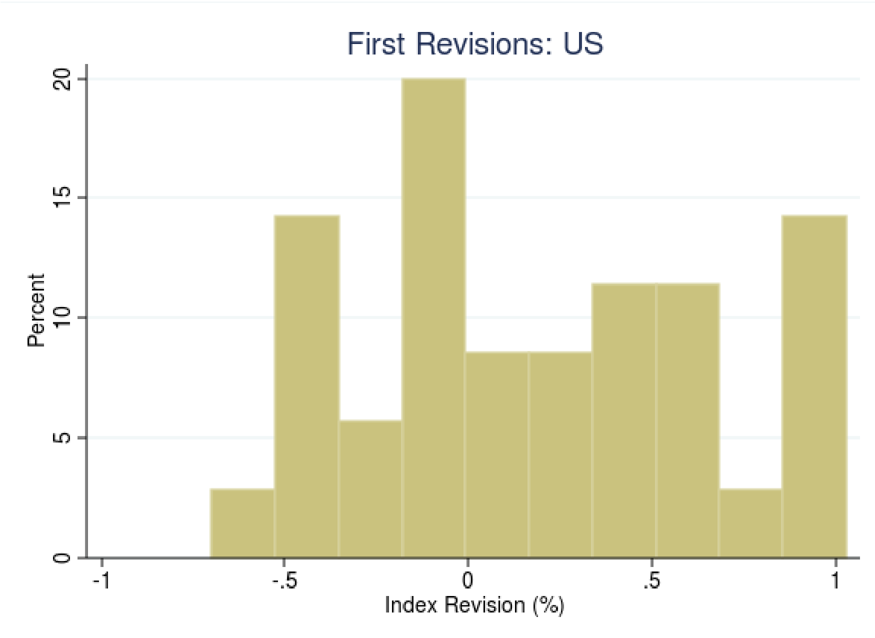
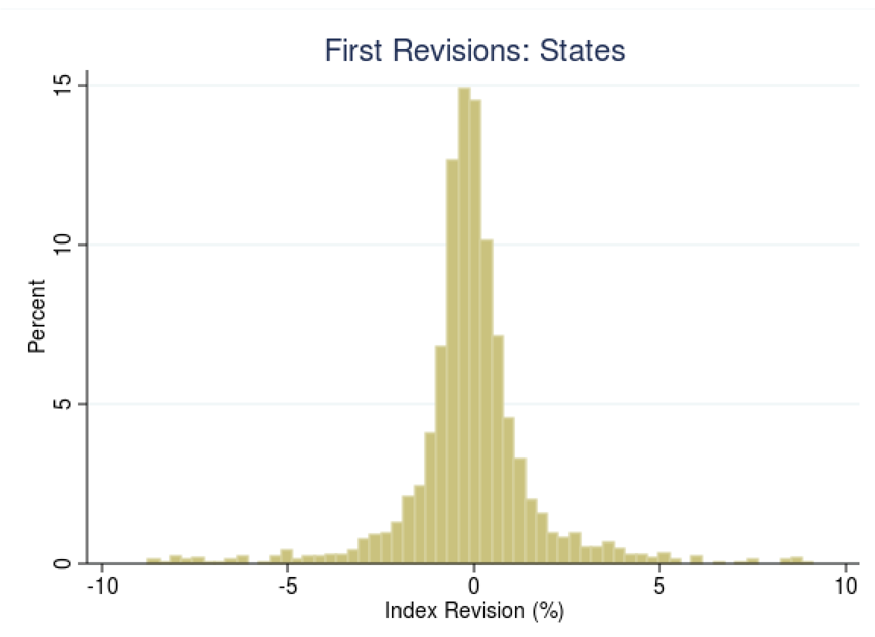


Figure 5
Average Cumulative Monthly Revisions
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
Average Cumulative Revision (%)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Cumulative Revision Number
Zip Codes
States
National

Figure 6: Histograms of the Distribution of First Revisions[9]





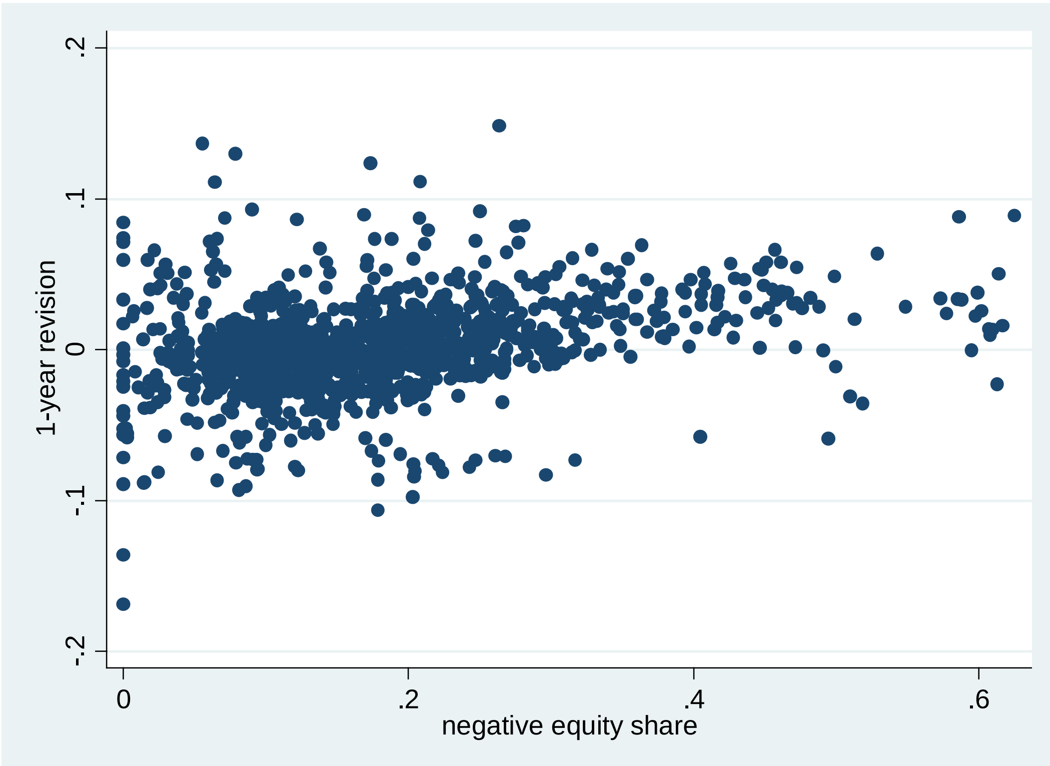
[9] First revisions (percent) to initial release of HPI for the national index and individual states: March 2011–February 2014; histogram for states truncated at ± 10%.

Table 2: Negative Equity, LTV, and Index Revisions[10]

Date	Revision to HPI	Negative Equity Share		Households Revised Out	Share w/LTV Above 80%	
	1st to 12th Release	1st Release	12th Release	(100,000's)	1st Release	12th Release
201103	1.20%	28.72%	27.60%	587	52.49%	51.72%
201104	1.43%	27.06%	25.73%	696	51.16%	50.43%
201105	1.15%	25.70%	24.59%	577	50.04%	49.51%
201106	1.34%	24.50%	23.23%	668	49.14%	48.42%
201107	1.00%	23.15%	22.32%	433	48.05%	47.77%
201108	1.00%	22.98%	22.02%	500	47.96%	47.57%
201109	1.65%	23.66%	22.08%	830	48.72%	47.78%
201110	1.77%	24.06%	22.49%	820	49.17%	48.29%
201111	1.53%	23.73%	22.51%	635	48.90%	48.28%
201112	1.72%	24.33%	22.99%	701	49.54%	48.76%
201201	1.51%	25.27%	24.21%	553	50.58%	50.09%
201202	1.57%	25.79%	24.59%	626	50.97%	50.38%
201203	1.51%	25.14%	23.56%	827	50.43%	49.56%
201204	1.06%	22.47%	21.55%	486	48.23%	47.83%
201205	0.77%	20.00%	19.46%	283	46.21%	46.03%
201206	0.72%	18.17%	17.82%	185	44.65%	44.46%
201207	0.50%	16.55%	16.50%	23	42.93%	43.17%
201208	0.47%	15.90%	15.91%	-8	42.38%	42.60%
201209	0.69%	15.78%	15.55%	119	42.37%	42.40%
201210	0.79%	15.60%	15.36%	128	42.14%	42.19%
201211	0.73%	15.57%	15.33%	125	42.31%	42.29%
201212	1.12%	15.60%	15.03%	301	42.34%	41.80%
201301	0.58%	15.21%	15.20%	6	41.77%	41.91%
201302	0.87%	15.38%	15.06%	166	41.50%	41.28%
201303	1.02%	14.76%	14.15%	316	40.32%	39.86%
Average	1.11%	21.00%	20.19%	423	46.57%	46.18%

10 The number of households revised out of negative equity = (Percentage point revision to negative equity shares) × (# of households in U.S.) × (Share of households that are homeowners) × (Fraction of homeowners with a mortgage) = (Percentage point revision to negative equity shares) × 115 m × 0.65 × 0.7.

Figure 7: Index Revisions and Negative Equity Share[11]



[11] The one-year HPI revision for each state-date pair (relative to the initial release), plotted against the share of households in that state-date with negative equity; initial release dates range from 201103 to 201304.

Table 3: Predictability of First Revision (m=1)

	Model														
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV
Intercept	-0.001	-0.001	-0.001	0.000	0.001	0.000	0.000	0.008	-0.002	0.007	0.004	-0.006	0.002	0.000	-0.001
	(-1.33)	(-1.71)	(-1.84)	-0.9	-1.4	(-0.22)	-0.080	-1.63	(-1.23)	-1.420	-1.220	(-2.88)*	-0.580	-0.040	(-0.61)
First Lagged Revision		-0.184	-0.229			-0.177	-0.236				-0.250	-0.325			-0.323
		(-3.69)*	(-4.10)*			(-3.58)*	(-4.25)*				(-6.31)*	(-6.21)*			(-6.2)*
Second Lagged Revision			0.055				0.055					-0.059			-0.066
			-1.120				-1.250					(-1.31)			(-1.54)
Third Lagged Revision			0.108				0.122					0.019			0.019
			(2.95)*				(-3.24)*					-0.550			-0.560
Fourth Lagged Revision			-0.003				0.010					-0.060			-0.058
			(-0.10)				-0.300					(-1.84)			(-1.96)
First Lagged HPI Apprec.				-0.258	-0.282	-0.188	-0.178						-0.412	-0.397	-0.321
				(-3.83)*	(-3.37)*	(-3.17)*	(-2.83)*						(-4.09)*	(-3.80)*	(-4.13)*
Second Lagged HPI Apprec.					0.097		0.148							-0.015	0.077
					-1.470		(-2.39)*							(-0.20)	-1.020
Third Lagged HPI Apprec.					-0.129		-0.162							-0.173	-0.172
					(-1.75)		(-2.19)*							(-2.29)*	(-2.83)*
Fourth Lagged HPI Apprec.					-0.038		-0.067							0.012	-0.025
					(-0.78)		(-1.46)							-0.250	(-0.48)
State F.E.	N	N	N	N	N	N	N	N	Y	Y	Y	Y	Y	Y	Y
Time F.E.	N	N	N	N	N	N	N	Y	N	Y	N	Y	Y	Y	Y
R^2		0.0371	0.0694	0.0392	0.0475	0.0582	0.0973	0.0338	0.0395	0.0733	0.1402	0.1776	0.1245	0.1322	0.212
N	1785	1734	1581	1785	1785	1768	1612	1785	1785	1785	1734	1581	1785	1785	1581

* Significant at 5% level. SE's clustered at state level. t-statistics reported below coefficients.

Table 4: Predictability of Longer-Term Revisions

	m=3							m=11						
	I	II	III	IV	V	VI	VII	I	II	III	IV	V	VI	VII
Intercept	-0.002	-0.002	-0.002	0.000	0.001	0.000	0.000	0.001	0.001	0.001	0.003	0.003	0.003	0.003
	(-1.60)	(-1.76)	(-1.70)	(-0.12)	(0.65)	(-.4)	(0.37)	-0.450	(-0.29)	-0.380	(1.09)	(1.16)	(1.03)	(1.08)
First Lagged Revision		-0.183	-0.168			-0.171	-0.158		-0.225	-0.210			-0.218	-0.183
		(-2.88)*	(-2.50)*			(-2.81)*	(-2.44)*		(-2.72)*	(-2.49)*			(-2.75)*	(-2.34)*
Second Lagged Revision			0.019				0.032			0.014				0.029
			(0.29)				(0.49)			(0.15)				(0.31)
Third Lagged Revision			0.005				0.029			0.103				0.117
			(0.07)				(0.45)			(1.5)				(1.83)
Fourth Lagged Revision			-0.057				-0.035			0.081				0.092
			(-1.20)				(-0.73)			(1.38)				(1.52)
First Lagged HPI Apprec.				-0.369	-0.332	-0.332	-0.241				-0.476	-0.395	-0.453	-0.309
				(-4.78)*	(-3.71)*	(-4.14)*	(-3.24)*				(-6.33)*	(-3.80)*	(-5.8)*	(-2.46)*
Second Lagged HPI Apprec.					-0.005		0.003					-0.056		-0.040
					(-0.07)		(0.04)					(-0.71)		(-0.45)
Third Lagged HPI Apprec.					-0.140		-0.153					-0.227		-0.282
					(-1.91)		(-1.88)					(-2.86)*		(-3.15)
Fourth Lagged HPI Apprec.					-0.135		-0.157					-0.048		0.043
					(-2.19)*		(-2.33)*					(-0.49)		(0.45)
State F.E.	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Time F.E.	N	N	N	N	N	N	N	N	N	N	N	N	N	N
R^2		0.0225	0.0224	0.0547	0.0772	0.0646	0.0829		0.0195	0.0186	0.0532	0.0692	0.0652	0.0708
N	1683	1632	1479	1683	1683	1664	1508	1275	1224	1071	1275	1275	1248	1092

* Significant at 5% level. SE's clustered at state level. t-statistics reported below coefficients.

Table 5: One-Month Revisions vs. Ex-Post Appreciation

Intercept	-0.001 (-1.55)	0.007 (-1.36)
Contemporaneous House Price Apprec.	0.048 (1.47)	0.054 (1.04)
State and Time F.E.	N	Y
R^2	0.0013	0.0741
N	1785	1785

References

Deng, Yongheng, and John M. Quigley. “Index Revision, House Price Risk, and the Market for House Price Derivatives.” *Journal of Real Estate and Financial Economics*, 37 (2008), pp. 191–209.

Dynan, Karen, and Douglas Elmendorf. “Do Provisional Estimates of Output Miss Economic Turning Points?” FEDS Working Paper (2001).

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt. “What ‘Triggers‘ Mortgage Default?” *American Economic Review*, 100 (2010), pp. 490–494.

Shiller, Robert J. “Arithmetic Repeat Sales Price Estimators,” *Journal of Housing Economics*, 1 (1991), pp. 110–126.

Tracy, Joseph, Richard Peach, and Joshua Abel. “First Impressions Can Be Misleading: Revisions to House Price Changes,” *Liberty Street Economics*, Federal Reserve Bank of New York, March 26, 2013.