Error Statistics for the Survey of Professional Forecasters for Housing Starts [QA, M.Units]

Release Date: 09/01/2009

Tom Stark
Assistant Director and Manager
Real-Time Data Research Center
Economic Research Department
Federal Reserve Bank of Philadelphia

1. OVERVIEW.

This document reports error statistics for median projections from the Survey of Professional Forecasters (SPF), conducted since 1990 by the Federal Reserve Bank of Philadelphia. We provide the results in a series of tables and, in the PDF version of this document, a number of charts. The tables show the survey variable forecast and, importantly, the transformation of the data that we used to generate the statistics. (The transformation is usually a quarter-over-quarter growth rate, expressed in annualized percentage points. However, some variables, such as interest rates, the unemployment rate, and housing starts are untransformed and, thus, expressed in their natural units.)

The paragraphs below explain the format of the tables and charts and the methods used to compute the statistics. These paragraphs are general. The same discussion applies to all variables in the survey.

2. DESCRIPTION OF TABLES.

Table 1 reports error statistics for various forecast horizons, sample periods, and choices of the real-time historical value that we used to assess accuracy. In each quarterly survey, we ask our panelists for their projections for the current quarter and the next four quarters. The current quarter is defined as the quarter in which we conducted the survey. Our tables provide error statistics separately for each quarter of this five-quarter horizon, beginning with the current quarter (denoted H = 1) and ending with the quarter that is four quarters in the future (H = 5). For each horizon, we report the mean forecast error [ME(S)], the mean absolute forecast error [MAE(S)], and the root-mean-square error [RMSE(S)]. All are standard measures of accuracy, though the academic literature generally places the most weight on the latter.

We define a forecast error as the difference between the historical value and the forecast. The mean error for each horizon is simply the average of the forecast errors at that horizon, constructed over the sample period shown in Table 1. Other things the same, a forecast with a mean error close to zero is better than one with a mean error far from zero. The mean absolute error is the sample average of the absolute value of the errors. Many analysts prefer this measure to the mean error because it does not allow large positive errors to offset large negative errors. In this sense, the mean absolute error gives a cleaner estimate of the size of the errors. Decision makers, however, may care not only about the average size of the errors but also about their variability, as measured by variance. Our last measure of accuracy is one that reflects the influence of the mean error and the variance of the error. The root-mean-square error for the SPF [RMSE(S)], the measure most often used by analysts and academicians, is the square root of the the average squared error. The lower the root-mean-square error, the more accurate the forecast.

2.1. Benchmark Models.

The forecast error statistics from the SPF are of interest in their own right. However, it is often more interesting to compare such statistics with those of alternative, or benchmark, forecasts. Table 1 reports four such comparisons. It shows the ratio of the root-mean-square error of the SPF forecast to that of four benchmark models. The benchmark models are statistical equations that we estimate on the data. We use the equations to generate projections for the same horizons included in the survey. In effect, we imagine standing back in time at each date when a survey was conducted and generating a separate forecast with each benchmark model. We do this in the same way that a survey panelist would have done using his own model.

A RMSE ratio below unity indicates that the SPF consensus (median) forecast has a root-mean-square error lower than that of the benchmark. This means the SPF is more accurate. We now describe the benchmark models. The first is perhaps the simplest of all possible benchmarks: A no-change model. In this model, the forecast for quarter T, the one-step-ahead or current-quarter forecast, is simply the historical value for the prior quarter (T - 1). There is, in other words, no change in the forecast compared with the historical value. Moreover, the forecast for the remaining quarters of the horizon is the same as the forecast for the current quarter. We denote the relative RMSE ratio for this benchmark as RMSE(S/NC), using NC to indicate no change. The second and third benchmark models generate projections using one or more historical observations of the the variable forecast, weighted by coefficients estimated from the data. Such autoregressive (AR) models can be formulated in two ways. We can estimate one model to generate the forecasts at all horizons, using an iteration method to generate the projections beyond the current quarter (IAR), or we can directly estimate a new model for each forecast horizon (DAR). The latter formulation has been shown to reduce the bias in a forecast when the underlying model is characterized by certain types of misspecification. The root-mean-square error ratios are denoted RMSE(S/IAR) and RMSE(S/DAR), respectively.

The one- through five-step-ahead projections of the benchmark models use information on the quarterly average of the variable forecast. The latest historical observation is for the quarter that is one quarter before the quarter of the first projection in the horizon. In contrast, the panelists generate their projections with the help of additional information. They submit their projections near the middle of each quarter and hence have access to some monthly indicators for the first month of each quarter, when those data are released before the survey deadline. This puts the projections of panelists for some variables at an advantage relative to the corresponding benchmark projections. Moreover, the panelists may also examine the very recent historical values of such monthly indicators in forming their projections for quarterly averages. Such monthly statistical momentum represents an advantage not shared by the benchmark models, which use only quarterly averages. For survey variables whose observations are reported at a monthly frequency, such as interest rates, industrial production, housing starts, and unemployment, we estimate and forecast a fourth benchmark model, the DARM. This model adds recent monthly historical values to the specification of the DAR model. For the projections for unemployment, nonfarm payroll employment, and interest rates, we add the values of monthly observations, beginning with that for the first month of the first quarter of the forecast horizon. These values should be in the information set of the survey panelists at the time they formed their projections. In contrast, for variables such as housing starts and industrial production, we include only lagged values of monthly observations. For such variables, the panelists would not have known the monthly observation for the first month of the first quarter of the forecast horizon. In general, we find that adding monthly observations to the benchmark DAR models improves accuracy. Indeed, for the projections for interest rates and the unemployment rate, the accuracy of the benchmark DARM projections rivals that of the SPF projections.

2.2. Real-Time Data.

All benchmark models are estimated on a rolling, fixed window of 60 real-time quarterly observations. Lag lengths, based on either the Akaike information criterion (AIC) or the Schwarz information criterion (SIC), are re-estimated each period. The tables below indicate whether the lag length was was chosen by the AIC or SIC.

We would like to make the comparison between the SPF forecast and the forecasts of each benchmark as fair as possible. Therefore, we must subject the benchmark models to the same data environment the survey panelists faced when they made their projections. This is important because macroeconomic data are revised often, and we do not want the benchmark models to use a data set that differs from the one our panelists would have used. We estimate and forecast the benchmark models with real-time data from the Philadelphia Fed real-time data set, using the vintage of data that the survey panelists would have had at the time they generated their own projections. (For more information on the Philadelphia Fed real-time data set, go to www.philadelphiafed.org/econ/forecast/real-time-data/.)

An open question in the literature on forecasting is: What version or vintage of the data should we use to compute the errors? A closely related question is: What version of the data are professional forecasters trying to predict? Our computations take no strong position on these questions. Rather, in Table 1, we evaluate the projections (SPF and benchmark) with five alternative measures of the historical values, all from the Philadelphia Fed real-time data set. These measures range from the initial-release values to the values as we know them today. All together, we compute the forecast error statistics using the following five alternative measures of historical values:

(1) The initial or first-release value;
(2) The revised value as it appears one quarter after the initial release;
(3) The revised value as it appears five quarters after the initial release;
(4) The revised value as it appears nine quarters after the initial release;
(5) The revised value as it appears today.

Each measure of the historical value has advantages and disadvantages. The initial-release value is the first measure released by government statistical agencies. A forecaster might be very interested in this measure because it enables him to evaluate his latest forecast soon after he generated it. However, early releases of the data are often subject to large measurement error. Subsequent releases [(2) - (5)] are more accurate, but they are available much later than the initial release. As we go from the first measure to the fifth, we get more reliability, at the cost of higher delays in availability.

The last two columns in Table 1 report the number of observations that we used to compute the error statistics. Some observations are omitted because the data are missing in the real-time data set, such as occurred when federal government statistical agencies closed in late 1995.

2.3. Recent Projections and Realizations.

Tables 2 to 7 provide information on recent projections and realizations. They show how we align the data prior to computing the forecast errors that form the backbone of the computations in Table 1. Any forecast error can be written as the equation given by error = realization - forecast. For our computations, we must be more precise because, for each projection (SPF and benchmarks), we have different periods forecast (T) different forecast horizons (h), and several measures of the realization (m). Thus, we can define the forecast error more precisely as

$$error(T, h, m) = realization(T, m) - forecast(T, h).$$

Tables 2 to 7 are organized along these lines. Table 2 shows recent forecasts from the SPF. Each column gives the projection for a different horizon or forecast step (h), beginning with that for the current quarter, defined as the quarter in which we conducted the survey. The dates (T) given in the rows show the periods forecast. These also correspond to the dates that we conducted the survey. Tables 3 to 6 report the recent projections of the four benchmark models. These are organized in the same way as Table 2. Table 7 reports recent values of the five alternative realizations (m) we use to compute the error statistics.

2.4. Qualifications.

We note two minor qualifications to the methods discussed above. The first concerns the vintage of data that we used to estimate and forecast the benchmark models for CPI inflation. The second concerns the five measures of realizations used for the unemployment rate, nonfarm payroll employment, and CPI inflation. To estimate and forecast the benchmark models for CPI inflation, we use the vintage of data that would have been available in the middle of each quarter. This predates by one month the vintage that SPF panelists would have had at their disposal when they formed their projections. The effect is likely small because revisions to the CPI are generally small. To compute the realizations for unemployment, nonfarm payroll employment, and CPI inflation, we use the vintages associated with the middle of each quarter. The measure that we call initial comes from this vintage, even though the initial estimate was available one month earlier. Thus, for these variables, our initial estimate reflects some revision by government statistical agencies. The effect for unemployment and CPI inflation is likely small. The effect could be somewhat larger for nonfarm payroll employment.

3. DESCRIPTION OF GRAPHS.

3.1. Root-Mean-Square Errors.

For each sample period shown in Table 1, we provide graphs of the root-mean-square error for the SPF forecast. There is one page for each sample period. On each page, we plot (for each forecast horizon) the RMSE on the y-axis. The x-axis shows the measure of the historical value that we used to compute the RMSE. These range from the value on its initial release to the value one quarter later to the value as we know it now (at the time we made the computation).

The graphs provide a tremendous amount of information. If we focus on a particular graph, we can see how a change in the measure of the realization (x-axis) affects the root-mean-square-error measure of accuracy. The effect is pronounced for some variables, such as real GDP and some of its components. For others, there is little or no effect. For example, because the historical data on interest rates are not revised, the estimated RMSE is the same in each case.

If we compare a particular point on one graph with the same point on another, we see how the forecast horizon affects accuracy. In general, the RMSE rises (accuracy falls) as the forecast horizon lengthens. Finally, if we compare a graph on one page with the corresponding graph on another page, we see how our estimates of accuracy in the SPF change with the sample period. Periods characterized by a high degree of economic turbulence will generally produce large RMSEs.

3.2. Fan Charts.

The last chart plots recent historical values and the latest SPF forecast. It also shows confidence intervals for the forecast, based on back-of-the-envelope calculations. The historical values and the SPF forecast are those associated with the latest vintage of data and survey, respectively, available at the time we ran our computer programs. The confidence intervals are constructed under the assumption that the historical forecast errors over the sample (shown in the footnote) follow a normal distribution with a mean of zero and a variance given by the root-mean-square error. The latter is estimated over the aforementioned sample, using the measure of history listed in the footnote.

Table 1. Forecast Error Statistics for SPF Variable: HOUSING (Housing Starts [QA, M.Units])

Computed Over Various Sample Periods
Various Measures of Realizations
Transformation: Level
Lag Length for IAR(p), DAR(p), and DARM(p) Models: AIC
Last Updated: 09/01/2009 10:22

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: Initial Release									
1985:01-2007:01									
1	0.02	0.07	0.08	0.82	0.80	0.80	0.82	89	89
2	0.03	0.10	0.12	0.88	0.83	0.84	0.79	89	89
3	0.03	0.12	0.15	0.92	0.85	0.84	0.85	89	89
4	0.04	0.14	0.17	0.98	0.89	0.88	0.89	89	89
5	0.04	0.16	0.20	0.99	0.90	0.86	0.89	89	89
1985:01-1996:04									
1	-0.01	0.07	0.08	0.72	0.70	0.70	0.73	48	48
2	-0.02	0.09	0.11	0.74	0.68	0.70	0.65	48	48
3	-0.03	0.10	0.12	0.76	0.68	0.70	0.71	48	48
4	-0.04	0.11	0.15	0.83	0.74	0.75	0.76	48	48
5	-0.06	0.13	0.18	0.85	0.77	0.74	0.77	48	48
1997:01-2007:01									
1	0.05	0.07	0.09	0.96	0.95	0.95	0.96	41	41
2	0.08	0.12	0.14	1.04	1.02	1.00	0.97	41	41
3	0.11	0.15	0.17	1.07	1.03	0.98	0.99	41	41
4	0.13	0.17	0.20	1.13	1.05	0.99	1.01	41	41
5	0.15	0.19	0.22	1.17	1.06	1.00	1.02	41	41

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: One Qtr After Initial Release									
1985:01-2007:01									
1	0.02	0.07	0.09	0.83	0.80	0.80	0.82	89	89
2	0.03	0.10	0.12	0.88	0.83	0.83	0.78	89	89
3	0.04	0.13	0.15	0.93	0.85	0.84	0.85	89	89
4	0.04	0.15	0.18	0.99	0.89	0.88	0.89	89	89
5	0.05	0.16	0.20	1.01	0.90	0.86	0.89	89	89
1985:01-1996:04									
1	-0.01	0.07	0.08	0.72	0.69	0.69	0.71	48	48
2	-0.02	0.09	0.11	0.74	0.67	0.69	0.63	48	48
3	-0.03	0.10	0.12	0.76	0.67	0.69	0.70	48	48
4	-0.04	0.12	0.15	0.83	0.74	0.75	0.76	48	48
5	-0.05	0.13	0.18	0.86	0.77	0.73	0.77	48	48
1997:01-2007:01									
1	0.05	0.08	0.09	0.98	0.97	0.97	0.96	41	41
2	0.08	0.12	0.14	1.05	1.03	1.01	0.97	41	41
3	0.11	0.16	0.18	1.09	1.04	0.99	1.00	41	41
4	0.14	0.18	0.21	1.14	1.06	1.00	1.01	41	41
5	0.16	0.20	0.22	1.18	1.06	1.01	1.03	41	41

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: Five Qtrs After Initial Release									
1985:01-2007:01									
1	0.02	0.07	0.08	0.82	0.79	0.79	0.79	89	89
2	0.03	0.10	0.12	0.88	0.82	0.83	0.77	89	89
3	0.03	0.12	0.15	0.92	0.85	0.84	0.85	89	89
4	0.04	0.14	0.18	0.97	0.88	0.87	0.88	89	89
5	0.04	0.16	0.20	1.00	0.90	0.85	0.88	89	89
1985:01-1996:04									
1	-0.01	0.06	0.08	0.71	0.67	0.67	0.68	48	48
2	-0.02	0.08	0.10	0.73	0.65	0.67	0.61	48	48
3	-0.03	0.10	0.12	0.75	0.66	0.68	0.70	48	48
4	-0.04	0.11	0.15	0.81	0.72	0.74	0.76	48	48
5	-0.06	0.13	0.17	0.85	0.76	0.72	0.76	48	48
1997:01-2007:01									
1	0.05	0.08	0.09	0.96	0.95	0.95	0.94	41	41
2	0.08	0.12	0.14	1.03	1.01	1.00	0.96	41	41
3	0.11	0.15	0.18	1.07	1.03	0.98	0.99	41	41
4	0.13	0.18	0.21	1.11	1.05	0.99	1.00	41	41
5	0.16	0.20	0.22	1.16	1.05	1.00	1.02	41	41

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: Nine Qtrs After Initial Release									
1985:01-2007:01									
1	0.01	0.07	0.08	0.81	0.77	0.77	0.78	89	89
2	0.02	0.10	0.12	0.87	0.81	0.82	0.76	89	89
3	0.03	0.12	0.15	0.91	0.84	0.83	0.85	89	89
4	0.04	0.14	0.17	0.96	0.87	0.86	0.88	89	89
5	0.04	0.16	0.20	1.00	0.90	0.85	0.88	89	89
1985:01-1996:04									
1	-0.01	0.06	0.07	0.70	0.65	0.65	0.66	48	48
2	-0.02	0.08	0.10	0.72	0.65	0.66	0.60	48	48
3	-0.03	0.09	0.12	0.74	0.66	0.67	0.70	48	48
4	-0.05	0.11	0.15	0.80	0.72	0.73	0.75	48	48
5	-0.06	0.13	0.17	0.86	0.76	0.72	0.76	48	48
1997:01-2007:01									
1	0.05	0.08	0.09	0.94	0.93	0.93	0.92	41	41
2	0.08	0.12	0.14	1.02	1.00	0.99	0.95	41	41
3	0.10	0.15	0.17	1.06	1.03	0.97	0.99	41	41
4	0.13	0.18	0.20	1.11	1.04	0.98	0.99	41	41
5	0.15	0.19	0.22	1.16	1.05	1.00	1.02	41	41

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: Latest Vintage									
1985:01-2007:01									
1	0.01	0.07	0.08	0.81	0.77	0.77	0.77	89	89
2	0.02	0.10	0.12	0.87	0.81	0.82	0.76	89	89
3	0.03	0.12	0.15	0.91	0.84	0.83	0.85	89	89
4	0.04	0.14	0.17	0.95	0.87	0.85	0.88	89	89
5	0.04	0.16	0.20	1.00	0.90	0.85	0.88	89	89
1985:01-1996:04									
1	-0.01	0.06	0.07	0.70	0.64	0.64	0.66	48	48
2	-0.02	0.08	0.10	0.72	0.64	0.66	0.60	48	48
3	-0.03	0.09	0.12	0.74	0.66	0.67	0.70	48	48
4	-0.05	0.11	0.15	0.79	0.71	0.72	0.75	48	48
5	-0.06	0.13	0.17	0.86	0.76	0.72	0.76	48	48
1997:01-2007:01									
1	0.04	0.08	0.09	0.94	0.93	0.93	0.92	41	41
2	0.08	0.12	0.14	1.01	1.00	0.99	0.95	41	41
3	0.10	0.15	0.17	1.06	1.03	0.97	0.98	41	41
4	0.13	0.18	0.20	1.10	1.04	0.98	0.99	41	41
5	0.15	0.19	0.22	1.16	1.05	1.00	1.02	41	41

Table 1 notes.

(1) The forecast horizon is given by H, where H = 1 is the SPF forecast for the current quarter.
(2) The headers ME(S), MAE(S), and RMSE(S) are mean error, mean absolute error, and root-mean-square error for the SPF.
(3) The header RMSE(S/NC) is the ratio of the SPF RMSE to that of the no-change (NC) model.
(4) The headers RMSE(S/IAR), RMSE(S/DAR) and RMSE(S/DARM) are the ratios of the SPF RMSE to the RMSE of the iterated and direct autoregressive models and the direct autoregressive model augmented with monthly observations, respectively. All models are estimated on a rolling window of 60 observations from the Phila Fed real-time data set.
(5) The headers Nspf and N are the number of observations analyzed for the SPF and benchmark models.
(6) When the variable forecast is a growth rate or an interest rate, it is expressed in annualized percentage points. When the variable forecast is the unemployment rate, it is expressed in percentage points.
(7) Sample periods refer to the dates forecast, not the dates when the forecasts were made.

Source: Tom Stark, Research Department, FRB Philadelphia.

Table 2. Recent SPF Forecasts - Dated at the Quarter Forecast

Variable: HOUSING (Housing Starts [QA, M.Units])
By Forecast Step (1 to 5)
Transformation: Level
Last Updated: 09/01/2009 10:22

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2002:04	1.660	1.620	1.586	1.580	1.554
2003:01	1.700	1.660	1.613	1.595	1.590
2003:02	1.700	1.667	1.645	1.616	1.595
2003:03	1.710	1.673	1.659	1.640	1.605
2003:04	1.800	1.685	1.640	1.650	1.658
2004:01	1.923	1.761	1.669	1.645	1.641
2004:02	1.920	1.867	1.730	1.660	1.650
2004:03	1.890	1.850	1.800	1.688	1.652
2004:04	1.925	1.825	1.787	1.790	1.662
2005:01	1.924	1.888	1.790	1.749	1.760
2005:02	1.990	1.894	1.850	1.763	1.740
2005:03	2.000	1.941	1.845	1.817	1.740
2005:04	2.008	1.969	1.926	1.811	1.800
2006:01	1.980	1.967	1.928	1.875	1.800
2006:02	1.965	1.936	1.928	1.883	1.839
2006:03	1.828	1.890	1.914	1.880	1.862
2006:04	1.658	1.800	1.840	1.880	1.844
2007:01	1.538	1.650	1.793	1.850	1.850
2007:02	1.464	1.547	1.625	1.767	1.809
2007:03	1.420	1.472	1.551	1.614	1.750
2007:04	1.200	1.400	1.494	1.583	1.635
2008:01	1.027	1.157	1.400	1.518	1.623
2008:02	0.956	1.015	1.150	1.425	1.529
2008:03	0.927	0.950	1.059	1.181	1.450
2008:04	0.810	0.925	0.964	1.090	1.235
2009:01	0.560	0.805	0.950	1.000	1.125
2009:02	0.521	0.562	0.820	0.977	1.069
2009:03	0.583	0.548	0.610	0.850	1.007
2009:04	NA	0.621	0.600	0.663	0.900
2010:01	NA	NA	0.690	0.662	0.720
2010:02	NA	NA	NA	0.722	0.743
2010:03	NA	NA	NA	NA	0.795

Table 2 notes.

(1) Each column gives the sequence of SPF projections for a given forecast step. The forecast steps range from one (the forecast for the quarter in which the survey was conducted) to four quarters in the future (step 5).
(2) The dates listed in the rows are the dates forecast, not the dates when the forecasts were made, with the exception of the forecast at step one, for which the two dates coincide.

Source: Tom Stark, Research Department, FRB Philadelphia.

Table 3. Recent Benchmark Model 1 IAR Forecasts - Dated at the Quarter Forecast

Variable: HOUSING (Housing Starts [QA, M.Units])
By Forecast Step (1 to 5)
Transformation: Level
Lag Length for IAR(p): AIC
Last Updated: 09/01/2009 10:22

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2002:04	1.687	1.623	1.639	1.516	1.518
2003:01	1.732	1.669	1.608	1.618	1.505
2003:02	1.741	1.717	1.652	1.594	1.599
2003:03	1.711	1.729	1.703	1.637	1.581
2003:04	1.859	1.700	1.719	1.690	1.622
2004:01	2.043	1.850	1.689	1.709	1.678
2004:02	1.939	2.045	1.842	1.679	1.699
2004:03	1.908	1.935	2.048	1.833	1.669
2004:04	1.968	1.905	1.931	2.051	1.826
2005:01	1.958	1.967	1.902	1.926	2.054
2005:02	2.093	1.957	1.967	1.899	1.922
2005:03	2.015	2.102	1.957	1.967	1.896
2005:04	2.074	2.018	2.110	1.956	1.967
2006:01	2.034	2.079	2.021	2.118	1.956
2006:02	2.126	2.033	2.083	2.024	2.126
2006:03	1.873	2.121	2.032	2.087	2.027
2006:04	1.736	1.869	2.117	2.032	2.092
2007:01	1.575	1.736	1.864	2.113	2.031
2007:02	1.487	1.585	1.736	1.860	2.109
2007:03	1.479	1.499	1.594	1.736	1.856
2007:04	1.321	1.494	1.510	1.602	1.737
2008:01	1.152	1.345	1.509	1.520	1.610
2008:02	1.042	1.180	1.366	1.522	1.530
2008:03	1.030	1.078	1.212	1.386	1.534
2008:04	0.832	1.056	1.117	1.242	1.405
2009:01	0.581	0.808	1.084	1.156	1.270
2009:02	0.449	0.535	0.815	1.111	1.191
2009:03	0.502	0.389	0.479	0.821	1.136
2009:04	NA	0.446	0.331	0.410	0.843
2010:01	NA	NA	0.413	0.273	0.372
2010:02	NA	NA	NA	0.425	0.212
2010:03	NA	NA	NA	NA	0.457

Table 3 notes.

(1) Each column gives the sequence of benchmark IAR projections for a given forecast step. The forecast steps range from one to five. The first step corresponds to the forecast that SPF panelists make for the quarter in which the survey is conducted.
(2) The dates listed in the rows are the dates forecast, not the dates when the forecasts were made, with the exception of the forecast at step one, for which the two dates coincide.
(3) The IAR benchmark model is estimated on a fixed 60-quarter rolling window. Its forecasts are computed with the indirect method. Estimation uses data from the Philadelphia Fed real-time data set.

Source: Tom Stark, Research Department, FRB Philadelphia.

Table 4. Recent Benchmark Model 2 No-Change Forecasts - Dated at the Quarter Forecast

Variable: HOUSING (Housing Starts [QA, M.Units])
By Forecast Step (1 to 5)
Transformation: Level
Last Updated: 09/01/2009 10:22

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2002:04	1.707	1.658	1.715	1.571	1.594
2003:01	1.747	1.707	1.658	1.715	1.571
2003:02	1.753	1.747	1.707	1.658	1.715
2003:03	1.723	1.753	1.747	1.707	1.658
2003:04	1.868	1.723	1.753	1.747	1.707
2004:01	2.040	1.868	1.723	1.753	1.747
2004:02	1.944	2.040	1.868	1.723	1.753
2004:03	1.912	1.944	2.040	1.868	1.723
2004:04	1.968	1.912	1.944	2.040	1.868
2005:01	1.959	1.968	1.912	1.944	2.040
2005:02	2.085	1.959	1.968	1.912	1.944
2005:03	2.012	2.085	1.959	1.968	1.912
2005:04	2.069	2.012	2.085	1.959	1.968
2006:01	2.035	2.069	2.012	2.085	1.959
2006:02	2.131	2.035	2.069	2.012	2.085
2006:03	1.878	2.131	2.035	2.069	2.012
2006:04	1.735	1.878	2.131	2.035	2.069
2007:01	1.564	1.735	1.878	2.131	2.035
2007:02	1.474	1.564	1.735	1.878	2.131
2007:03	1.462	1.474	1.564	1.735	1.878
2007:04	1.296	1.462	1.474	1.564	1.735
2008:01	1.151	1.296	1.462	1.474	1.564
2008:02	1.035	1.151	1.296	1.462	1.474
2008:03	1.016	1.035	1.151	1.296	1.462
2008:04	0.879	1.016	1.035	1.151	1.296
2009:01	0.656	0.879	1.016	1.035	1.151
2009:02	0.523	0.656	0.879	1.016	1.035
2009:03	0.541	0.523	0.656	0.879	1.016
2009:04	NA	0.541	0.523	0.656	0.879
2010:01	NA	NA	0.541	0.523	0.656
2010:02	NA	NA	NA	0.541	0.523
2010:03	NA	NA	NA	NA	0.541

Table 4 notes.

(1) Each column gives the sequence of benchmark no-change projections for a given forecast step. The forecast steps range from one to five. The first step corresponds to the forecast that SPF panelists make for the quarter in which the survey is conducted.
(2) The dates listed in the rows are the dates forecast, not the dates when the forecasts were made, with the exception of the forecast at step one, for which the two dates coincide.
(3) The projections use data from the Philadelphia Fed real-time data set.

Source: Tom Stark, Research Department, FRB Philadelphia.

Table 5. Recent Benchmark Model 3 DAR Forecasts - Dated at the Quarter Forecast

Variable: HOUSING (Housing Starts [QA, M.Units])
By Forecast Step (1 to 5)
Transformation: Level
Lag Length for DAR(p): AIC
Last Updated: 09/01/2009 10:22

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2002:04	1.687	1.620	1.633	1.509	1.514
2003:01	1.732	1.668	1.598	1.603	1.499
2003:02	1.741	1.710	1.643	1.576	1.581
2003:03	1.711	1.728	1.689	1.617	1.557
2003:04	1.859	1.701	1.707	1.657	1.598
2004:01	2.043	1.842	1.689	1.684	1.637
2004:02	1.939	2.027	1.825	1.669	1.660
2004:03	1.908	1.942	2.004	1.803	1.655
2004:04	1.968	1.905	1.933	1.975	1.783
2005:01	1.958	1.966	1.910	1.910	1.952
2005:02	2.093	1.965	1.964	1.904	1.895
2005:03	2.015	2.089	1.961	1.964	1.876
2005:04	2.074	2.036	2.096	1.956	1.958
2006:01	2.034	2.082	2.030	2.092	1.960
2006:02	2.126	2.046	2.092	2.053	2.091
2006:03	1.873	2.138	2.051	2.098	2.035
2006:04	1.736	1.883	2.142	2.063	2.102
2007:01	1.575	1.740	1.958	2.161	2.073
2007:02	1.487	1.587	1.753	1.908	2.172
2007:03	1.479	1.505	1.599	1.765	1.958
2007:04	1.321	1.493	1.523	1.615	1.779
2008:01	1.152	1.309	1.512	1.540	1.628
2008:02	1.042	1.129	1.334	1.530	1.556
2008:03	1.030	1.012	1.159	1.365	1.545
2008:04	0.832	0.984	0.980	1.182	1.406
2009:01	0.581	0.821	0.989	1.035	1.243
2009:02	0.449	0.574	0.822	0.993	1.078
2009:03	0.502	0.407	0.568	0.850	1.090
2009:04	NA	0.431	0.382	0.594	0.895
2010:01	NA	NA	0.398	0.420	0.647
2010:02	NA	NA	NA	0.430	0.502
2010:03	NA	NA	NA	NA	0.565

Table 5 notes.

(1) Each column gives the sequence of benchmark DAR projections for a given forecast step. The forecast steps range from one to five. The first step corresponds to the forecast that SPF panelists make for the quarter in which the survey is conducted.
(2) The dates listed in the rows are the dates forecast, not the dates when the forecasts were made, with the exception of the forecast at step one, for which the two dates coincide.
(3) The DAR benchmark model is estimated on a fixed 60-quarter rolling window. Its forecasts are computed with the direct method. Estimation uses data from the Philadelphia Fed real-time data set.

Source: Tom Stark, Research Department, FRB Philadelphia.

Table 6. Recent Benchmark Model 4 DARM Forecasts - Dated at the Quarter Forecast

Variable: HOUSING (Housing Starts [QA, M.Units])
By Forecast Step (1 to 5)
Transformation: Level
Lag Length for DARM(p): AIC
Last Updated: 09/01/2009 10:22

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2002:04	1.743	1.637	1.603	1.558	1.446
2003:01	1.786	1.722	1.647	1.609	1.563
2003:02	1.716	1.756	1.753	1.655	1.616
2003:03	1.761	1.716	1.783	1.680	1.670
2003:04	1.849	1.746	1.666	1.747	1.637
2004:01	2.066	1.844	1.782	1.623	1.718
2004:02	1.952	2.050	1.810	1.754	1.582
2004:03	1.847	1.960	2.044	1.789	1.749
2004:04	1.941	1.829	1.962	2.033	1.751
2005:01	1.950	1.908	1.815	1.942	2.028
2005:02	1.950	1.976	1.898	1.812	1.894
2005:03	2.009	1.954	1.971	1.924	1.829
2005:04	2.099	2.019	1.893	1.898	1.923
2006:01	2.000	2.097	1.999	1.846	1.837
2006:02	2.096	2.038	2.136	2.027	1.907
2006:03	1.870	2.141	2.003	2.145	1.995
2006:04	1.751	1.914	2.136	1.997	2.139
2007:01	1.601	1.766	1.916	2.136	2.034
2007:02	1.498	1.601	1.790	1.895	2.053
2007:03	1.477	1.509	1.617	1.779	1.950
2007:04	1.251	1.500	1.516	1.597	1.916
2008:01	1.079	1.266	1.528	1.542	1.621
2008:02	1.002	1.072	1.303	1.547	1.537
2008:03	1.047	0.972	1.119	1.323	1.557
2008:04	0.801	1.002	0.956	1.185	1.358
2009:01	0.536	0.790	1.014	1.027	1.225
2009:02	0.477	0.535	0.815	1.011	1.070
2009:03	0.524	0.410	0.541	0.869	1.121
2009:04	NA	0.447	0.375	0.585	0.884
2010:01	NA	NA	0.401	0.415	0.607
2010:02	NA	NA	NA	0.424	0.506
2010:03	NA	NA	NA	NA	0.576

Table 6 notes.

(1) Each column gives the sequence of benchmark DARM projections for a given forecast step. The forecast steps range from one to five. The first step corresponds to the forecast that SPF panelists make for the quarter in which the survey is conducted.
(2) The dates listed in the rows are the dates forecast, not the dates when the forecasts were made, with the exception of the forecast at step one, for which the two dates coincide.
(3) The DARM benchmark model is estimated on a fixed 60-quarter rolling window. Its forecasts are computed with the direct method and incorporate recent monthly values of the dependent variable. Estimation uses data from the Philadelphia Fed real-time data set.

Source: Tom Stark, Research Department, FRB Philadelphia.

Table 7. Recent Realizations (Various Measures)
Source: Philadelphia Fed Real-Time Data Set

Variable: HOUSING (Housing Starts [QA, M.Units])
Transformation: Level
Last Updated: 09/01/2009 10:22

1- Initial Release
2- One Qtr After Initial Release
3- Five Qtrs After Initial Release
4- Nine Qtrs After Initial Release
5- Latest Vintage

Obs. Date	(1)	(2)	(3)	(4)	(5)
2002:04	1.747	1.748	1.743	1.730	1.730
2003:01	1.753	1.737	1.747	1.736	1.736
2003:02	1.723	1.739	1.745	1.754	1.754
2003:03	1.868	1.884	1.883	1.890	1.890
2003:04	2.040	2.031	2.035	2.036	2.036
2004:01	1.944	1.943	1.929	1.918	1.918
2004:02	1.912	1.920	1.923	1.937	1.937
2004:03	1.968	1.969	1.974	1.977	1.977
2004:04	1.959	1.975	1.973	1.965	1.965
2005:01	2.085	2.083	2.069	2.072	2.072
2005:02	2.012	2.044	2.064	2.051	2.051
2005:03	2.069	2.101	2.101	2.100	2.100
2005:04	2.035	2.059	2.060	2.069	2.069
2006:01	2.131	2.123	2.127	2.120	2.120
2006:02	1.878	1.873	1.861	1.855	1.855
2006:03	1.735	1.714	1.704	1.702	1.702
2006:04	1.564	1.559	1.555	1.570	1.570
2007:01	1.474	1.460	1.453	1.461	1.461
2007:02	1.462	1.464	1.460	NA	1.451
2007:03	1.296	1.300	1.298	NA	1.289
2007:04	1.151	1.151	1.151	NA	1.166
2008:01	1.035	1.053	1.059	NA	1.059
2008:02	1.016	1.025	NA	NA	1.017
2008:03	0.879	0.876	NA	NA	0.868
2008:04	0.656	0.660	NA	NA	0.658
2009:01	0.523	0.528	NA	NA	0.528
2009:02	0.541	NA	NA	NA	0.541
2009:03	NA	NA	NA	NA	NA
2009:04	NA	NA	NA	NA	NA
2010:01	NA	NA	NA	NA	NA
2010:02	NA	NA	NA	NA	NA
2010:03	NA	NA	NA	NA	NA

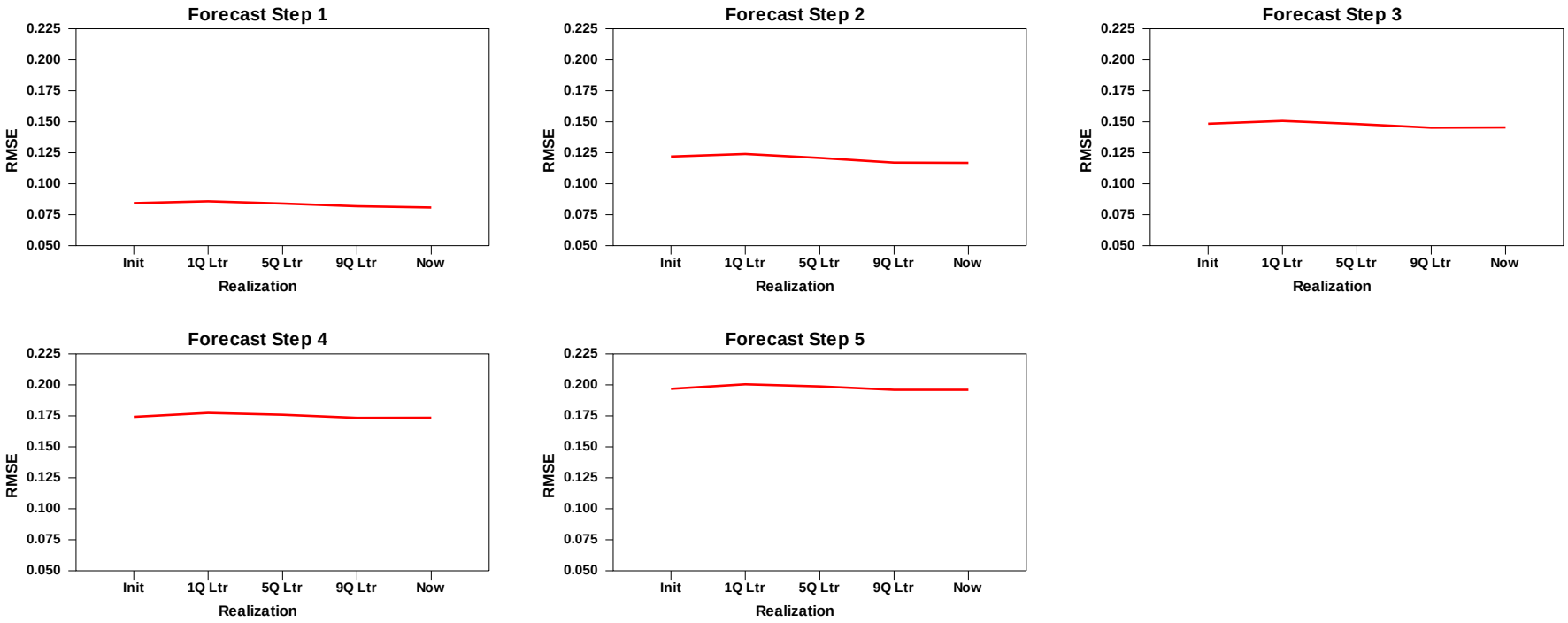
Table 7 notes.

(1) Each column reports a sequence of realizations from the Philadelphia Fed real-time data set.
(2) The date listed in each row is the observation date.
(3) Moving across a particular row shows how the observation is revised in subsequent releases.

Source: Tom Stark, Research Department, FRB Philadelphia.

Root-Mean-Square Errors: 1985:01-2007:01

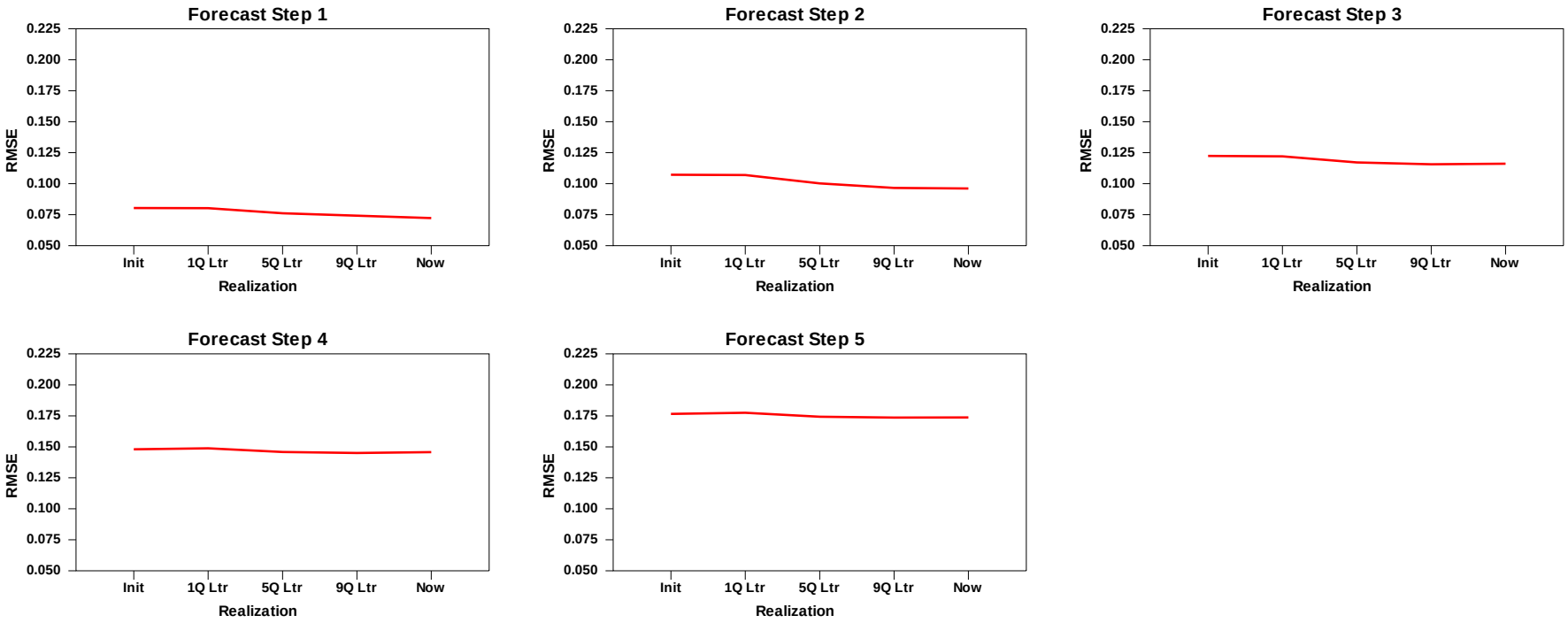
SPF Projections for Housing Starts [QA, M.Units], Transformation: Level



The RMSE is plotted against the realization used to compute it, from the value on initial release to the value as we now know it. Source: Tom Stark, FRB Philadelphia.

Root-Mean-Square Errors: 1985:01-1996:04

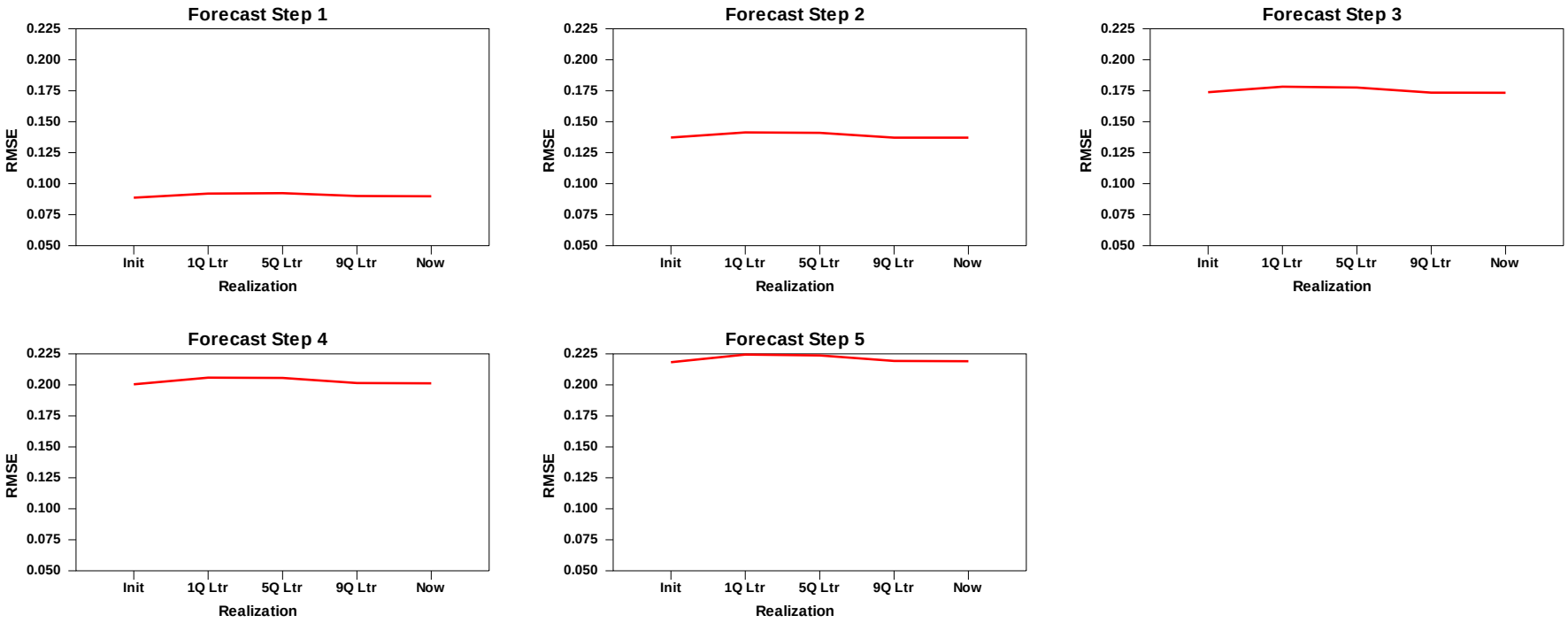
SPF Projections for Housing Starts [QA, M.Units], Transformation: Level



The RMSE is plotted against the realization used to compute it, from the value on initial release to the value as we now know it. Source: Tom Stark, FRB Philadelphia.

Root-Mean-Square Errors: 1997:01-2007:01

SPF Projections for Housing Starts [QA, M.Units], Transformation: Level



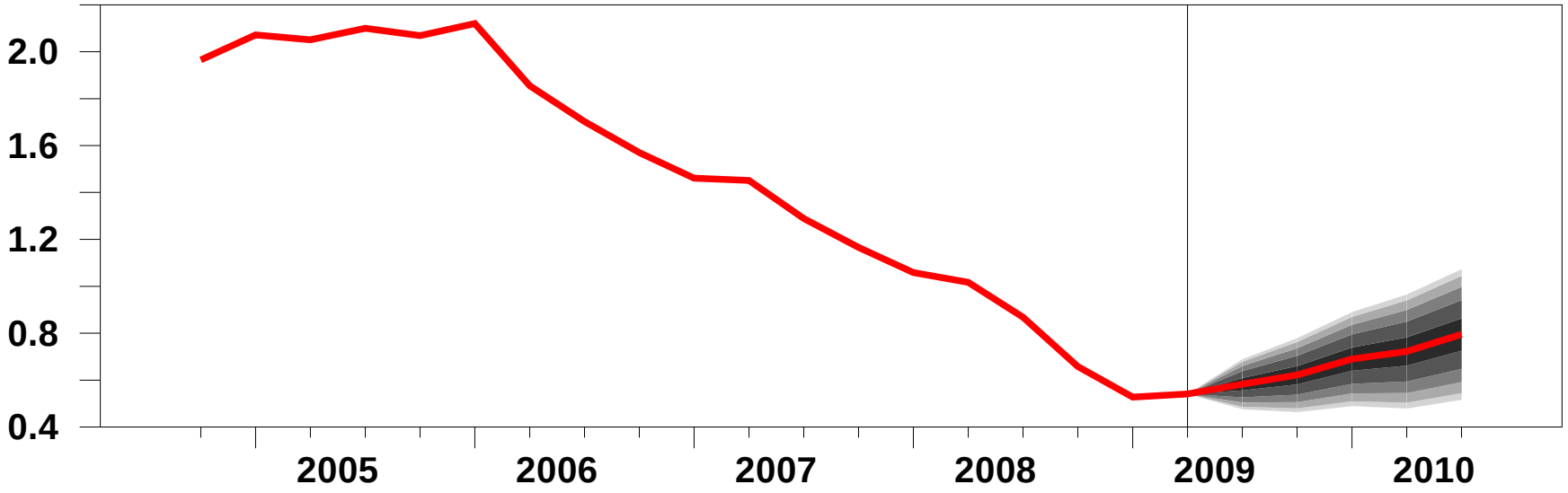
The RMSE is plotted against the realization used to compute it, from the value on initial release to the value as we now know it. Source: Tom Stark, FRB Philadelphia.

Housing Starts [QA, M.Units]

History, Forecasts, and Ranges for the SPF of 2009:03

Level

(Ranges Cover 25 to 80 Percent Confidence)



Ranges at each horizon use the N(0,MSE) density. The MSEs are based on the sample 85:01-08:01 and use the realization: Five Qtrs After Initial Release. Source: T.Stark, FRB Phila.