

Error Statistics for the Survey of Professional Forecasters: Housing Starts

Release Date: 08/21/2023

Tom Stark
 Assistant Vice President | Assistant Director
 Real-Time Data Research Center
 Economic Research Department
 Federal Reserve Bank of Philadelphia

Source for Historical Realizations: Census Bureau via Haver Analytics

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.

Tables 1A-1B report 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 periods shown in Table 1A. 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 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. Tables 1A-1B report four such comparisons. They show 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.

Table 1A reports the root-mean-square-error ratios using as many observations as possible for each model. The number of observations can differ from model to model. We first compute the RMSE for each model. We then construct the ratio.

Table 1B reports RMSE ratios after we adjust the samples to include only the observations common to both models in the pair. Accordingly, the ratios reported in Table 1B may differ slightly from those of Table 1A, depending on the availability of sufficient real-time observations for estimating the benchmark models or for computing the errors of the SPF or benchmark forecasts. Table 1B also reports three two-sided p-values for each ratio. The p-values, corrected for the presence of heteroskedasticity and serial correlation in the time series of differences in squared forecast errors, are those for the test of equality of mean-square error between the SPF and the benchmark. The p-values are those for:

- (1) The Diebold-Mariano statistic (July 1995, *Journal of Business and Economic Statistics*), using a uniform lag window with the truncation lag set to the forecast horizon minus unity. When the uniform lag window produces a nonpositive standard error, the Bartlett window is used.
- (2) The Harvey-Leybourne-Newbold correction (1997, *International Journal of Forecasting*) to the Diebold-Mariano statistic.
- (3) The Diebold-Mariano statistic, using a Bartlett lag window with the truncation lag increased four quarters beyond that of (1) and (2).

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 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 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. In Tables 1A - 1B, 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 1A 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 Tables 1A - 1B. Any error can be written as the equation given by $\text{error} = \text{realization} - \text{forecast}$. 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

$$\text{error}(T, h, m) = \text{realization}(T, m) - \text{forecast}(T, h).$$

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. They 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 postdates by one month the vintage that SPF panelists would have had at their disposal when they formed their projections.

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 in the vintage dated 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 squared root-mean-square error. The latter is estimated over the aforementioned sample, using the measure of history listed in the footnote.

4. SPECIAL TREATMENT OF COVID-19 EXTREME HISTORICAL OBSERVATIONS.

Many macroeconomic variables experienced extreme values in 2020:Q2 and 2020:Q3 due to the partial shutdown of the U.S. economy at the end of March 2020. In some cases, these extreme values adversely affected the estimation of parameters in the benchmark models. The effect produced unrealistic parameter values in the models. In some cases, the models became dynamically unstable in sample periods that encompassed the 2020:Q2 and 2020:Q3 observations, leading to distorted forecast error statistics in those benchmark models for some survey variables. Comparisons with the SPF projections became hard to defend.

Beginning with the error statistics published following the 2022:Q1 survey, we proceed in the following way. For some survey variables, we scale back the magnitude of the historical extreme observations for 2020:Q2 and 2020:Q3. We make these adjustment before the estimation of parameters in the benchmark models but reverse the adjustment prior to forecasting the models. The intent is to prevent adverse, unrealistic effects on the parameter estimates while allowing the extreme historical observations to condition the projections. These adjustments to extreme historical observations do not change the historical realizations used to compute forecast errors or forecast error statistics: We continue to use unadjusted, official U.S. government data for these purposes.

The survey variables for which we make the adjustments are: Nominal GDP, Unemployment, Employment, Industrial Production, Real GDP, and Real Personal Consumption Expenditures. We do not adjust the historical values for the remaining variables.

Table 1A.
Forecast Error Summary Statistics for SPF Variable: HOUSING (Housing Starts)

Computed Over Various Sample Periods

Various Measures of Realizations

Transformation: Level

Lag Length for IAR(p), DAR(p), and DARM(p) Models: AIC

Source for Historical Realizations: Census Bureau via Haver Analytics

Last Updated: 08/21/2023 13:25

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: Initial Release									
1985:01-2021:01									
1	0.01	0.06	0.08	0.76	0.69	0.69	0.76	145	145
2	0.01	0.10	0.13	0.88	0.77	0.82	0.78	145	145
3	-0.00	0.12	0.16	0.91	0.79	0.82	0.82	145	145
4	-0.02	0.15	0.19	0.97	0.83	0.85	0.86	145	145
5	-0.03	0.18	0.22	0.97	0.83	0.81	0.82	145	145
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
1997:01-2021:01									
1	0.02	0.06	0.08	0.78	0.69	0.69	0.78	97	97
2	0.02	0.11	0.14	0.94	0.80	0.87	0.85	97	97
3	0.01	0.14	0.17	0.97	0.82	0.86	0.86	97	97
4	-0.00	0.17	0.21	1.03	0.86	0.88	0.89	97	97
5	-0.02	0.20	0.24	1.01	0.84	0.83	0.83	97	97
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									
1985:01-2021:01									
1	0.01	0.07	0.08	0.78	0.71	0.71	0.77	145	145
2	0.01	0.10	0.13	0.88	0.77	0.82	0.78	145	145
3	0.00	0.13	0.16	0.92	0.79	0.82	0.82	145	145
4	-0.01	0.15	0.19	0.97	0.83	0.85	0.85	145	145
5	-0.03	0.18	0.22	0.97	0.83	0.81	0.82	145	145
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
1997:01-2021:01									
1	0.02	0.07	0.08	0.81	0.72	0.72	0.80	97	97
2	0.02	0.11	0.14	0.95	0.80	0.88	0.85	97	97
3	0.01	0.14	0.17	0.97	0.83	0.86	0.86	97	97
4	0.00	0.17	0.21	1.02	0.86	0.89	0.89	97	97
5	-0.01	0.20	0.24	1.01	0.84	0.83	0.83	97	97

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									
1985:01-2021:01									
	1	0.01	0.07	0.08	0.77	0.70	0.70	0.76	145 145
	2	0.01	0.10	0.13	0.88	0.76	0.82	0.78	145 145
	3	-0.00	0.12	0.15	0.91	0.78	0.81	0.82	145 145
	4	-0.01	0.15	0.19	0.96	0.82	0.84	0.85	145 145
	5	-0.03	0.18	0.22	0.97	0.82	0.81	0.82	145 145
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
1997:01-2021:01									
	1	0.02	0.07	0.08	0.80	0.71	0.71	0.80	97 97
	2	0.02	0.11	0.14	0.94	0.80	0.87	0.85	97 97
	3	0.01	0.14	0.17	0.97	0.82	0.86	0.86	97 97
	4	0.00	0.17	0.21	1.01	0.86	0.88	0.88	97 97
	5	-0.02	0.20	0.24	1.01	0.84	0.83	0.83	97 97
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									
1985:01-2021:01									
	1	0.01	0.07	0.08	0.76	0.69	0.69	0.74	145 145
	2	0.01	0.10	0.12	0.88	0.76	0.81	0.77	145 145
	3	-0.00	0.12	0.15	0.91	0.78	0.81	0.82	145 145
	4	-0.02	0.15	0.19	0.96	0.82	0.84	0.85	145 145
	5	-0.03	0.17	0.22	0.97	0.82	0.80	0.82	145 145
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
1997:01-2021:01									
	1	0.02	0.07	0.08	0.79	0.71	0.71	0.79	97 97
	2	0.02	0.11	0.13	0.93	0.80	0.87	0.85	97 97
	3	0.01	0.14	0.17	0.96	0.82	0.86	0.86	97 97
	4	-0.00	0.17	0.20	1.01	0.85	0.88	0.88	97 97
	5	-0.02	0.19	0.24	1.00	0.84	0.83	0.83	97 97

H	ME(S)	MAE(S)	RMSE(S)	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	Nspf	N
History: Latest Vintage									
1985:01-2021:01									
1	0.01	0.06	0.08	0.76	0.68	0.68	0.74	145	145
2	0.01	0.10	0.12	0.88	0.76	0.81	0.77	145	145
3	-0.00	0.12	0.15	0.90	0.78	0.81	0.82	145	145
4	-0.02	0.15	0.19	0.95	0.82	0.84	0.85	145	145
5	-0.03	0.17	0.22	0.97	0.82	0.80	0.81	145	145
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
1997:01-2021:01									
1	0.02	0.07	0.08	0.78	0.70	0.70	0.78	97	97
2	0.02	0.11	0.14	0.94	0.80	0.87	0.85	97	97
3	0.01	0.14	0.17	0.97	0.82	0.86	0.86	97	97
4	-0.00	0.17	0.20	1.01	0.85	0.88	0.88	97	97
5	-0.02	0.19	0.24	1.01	0.84	0.83	0.83	97	97

Notes for Table 1A.

- (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 Philadelphia 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 1B.
Ratios of Root-Mean-Square Errors for SPF Variable: HOUSING (Housing Starts)
Alternative P-Values in Parentheses

Computed Over Various Sample Periods
Various Measures of Realizations
Transformation: Level
Lag Length for IAR(p), DAR(p), and DARM(p) Models: AIC
Source for Historical Realizations: Census Bureau via Haver Analytics
Last Updated: 08/21/2023 13:25

History: Initial Release 1985:01-2021:01									
H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4	
1	0.757 (0.002) (0.003) (0.008)	0.693 (0.001) (0.001) (0.018)	0.693 (0.001) (0.001) (0.018)	0.764 (0.002) (0.002) (0.006)	145	145	145	142	
2	0.879 (0.015) (0.017) (0.052)	0.766 (0.001) (0.002) (0.004)	0.822 (0.003) (0.004) (0.008)	0.790 (0.001) (0.001) (0.001)	145	145	145	142	
3	0.912 (0.233) (0.243) (0.238)	0.786 (0.005) (0.006) (0.008)	0.817 (0.032) (0.037) (0.031)	0.829 (0.039) (0.045) (0.035)	145	145	145	142	
4	0.974 (0.753) (0.759) (0.735)	0.831 (0.048) (0.055) (0.046)	0.852 (0.114) (0.125) (0.091)	0.860 (0.092) (0.103) (0.077)	145	145	145	142	
5	0.971 (0.724) (0.732) (0.690)	0.827 (0.027) (0.034) (0.022)	0.809 (0.031) (0.039) (0.020)	0.825 (0.046) (0.055) (0.028)	145	145	145	142	

History: Initial Release 1985:01-1996:04									
H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4	
1	0.721 (0.000) (0.001) (0.000)	0.699 (0.002) (0.003) (0.000)	0.699 (0.002) (0.003) (0.000)	0.729 (0.028) (0.035) (0.014)	48	48	48	47	
2	0.741 (0.000) (0.000) (0.000)	0.680 (0.006) (0.011) (0.005)	0.700 (0.015) (0.023) (0.014)	0.651 (0.014) (0.022) (0.001)	48	48	48	47	
3	0.757 (0.000) (0.000) (0.000)	0.679 (0.021) (0.033) (0.018)	0.695 (0.043) (0.061) (0.033)	0.712 (0.054) (0.074) (0.023)	48	48	48	47	
4	0.826 (0.001) (0.005) (0.007)	0.738 (0.038) (0.061) (0.034)	0.753 (0.081) (0.113) (0.059)	0.768 (0.024) (0.042) (0.012)	48	48	48	47	
5	0.851 (0.127) (0.173) (0.099)	0.773 (0.058) (0.092) (0.049)	0.736 (0.053) (0.086) (0.038)	0.773 (0.068) (0.105) (0.043)	48	48	48	48	

History: Initial Release
1997:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.777 (0.053) (0.057) (0.104)	0.691 (0.018) (0.021) (0.100)	0.691 (0.018) (0.021) (0.100)	0.784 (0.021) (0.024) (0.062)	97	97	97	95
2	0.940 (0.353) (0.363) (0.469)	0.800 (0.024) (0.028) (0.049)	0.873 (0.055) (0.062) (0.103)	0.852 (0.013) (0.016) (0.062)	97	97	97	95
3	0.970 (0.754) (0.760) (0.756)	0.822 (0.048) (0.057) (0.068)	0.859 (0.169) (0.184) (0.170)	0.869 (0.185) (0.200) (0.189)	97	97	97	95
4	1.026 (0.813) (0.820) (0.796)	0.861 (0.187) (0.206) (0.184)	0.884 (0.315) (0.335) (0.283)	0.890 (0.293) (0.314) (0.272)	97	97	97	95
5	1.012 (0.904) (0.909) (0.891)	0.842 (0.097) (0.117) (0.085)	0.832 (0.120) (0.141) (0.090)	0.841 (0.137) (0.161) (0.102)	97	97	97	94

History: One Qtr After Initial Release
1985:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.775 (0.001) (0.002) (0.006)	0.707 (0.001) (0.001) (0.013)	0.707 (0.001) (0.001) (0.013)	0.772 (0.001) (0.001) (0.002)	145	145	145	142
2	0.882 (0.023) (0.026) (0.065)	0.766 (0.001) (0.002) (0.003)	0.821 (0.004) (0.005) (0.009)	0.787 (0.001) (0.001) (0.002)	145	145	145	142
3	0.918 (0.277) (0.287) (0.280)	0.788 (0.006) (0.008) (0.009)	0.818 (0.037) (0.042) (0.035)	0.829 (0.042) (0.047) (0.037)	145	145	145	142
4	0.973 (0.756) (0.762) (0.738)	0.830 (0.048) (0.056) (0.045)	0.851 (0.112) (0.123) (0.088)	0.860 (0.094) (0.104) (0.078)	145	145	145	142
5	0.974 (0.751) (0.759) (0.720)	0.826 (0.029) (0.036) (0.023)	0.810 (0.033) (0.040) (0.020)	0.824 (0.046) (0.055) (0.028)	145	145	145	142

History: One Qtr After Initial Release
1985:01-1996:04

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.719 (0.000) (0.001) (0.000)	0.687 (0.001) (0.003) (0.000)	0.687 (0.001) (0.003) (0.000)	0.709 (0.018) (0.024) (0.005)	48	48	48	47
2	0.737 (0.000) (0.000) (0.000)	0.668 (0.005) (0.009) (0.004)	0.687 (0.011) (0.017) (0.010)	0.635 (0.012) (0.019) (0.001)	48	48	48	47
3	0.761 (0.000) (0.000) (0.000)	0.674 (0.020) (0.032) (0.016)	0.690 (0.041) (0.059) (0.030)	0.702 (0.052) (0.072) (0.021)	48	48	48	47
4	0.830 (0.002) (0.006) (0.009)	0.735 (0.035) (0.056) (0.029)	0.749 (0.074) (0.104) (0.052)	0.767 (0.025) (0.044) (0.013)	48	48	48	47
5	0.858 (0.160) (0.209) (0.129)	0.769 (0.051) (0.083) (0.042)	0.732 (0.046) (0.077) (0.032)	0.767 (0.061) (0.096) (0.037)	48	48	48	48

History: One Qtr After Initial Release
1997:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.806 (0.055) (0.059) (0.109)	0.717 (0.021) (0.024) (0.100)	0.717 (0.021) (0.024) (0.100)	0.807 (0.014) (0.017) (0.047)	97	97	97	95
2	0.945 (0.428) (0.437) (0.525)	0.805 (0.027) (0.032) (0.053)	0.879 (0.079) (0.087) (0.129)	0.857 (0.023) (0.028) (0.078)	97	97	97	95
3	0.975 (0.802) (0.807) (0.802)	0.827 (0.058) (0.068) (0.080)	0.864 (0.193) (0.208) (0.191)	0.873 (0.200) (0.215) (0.204)	97	97	97	95
4	1.024 (0.831) (0.837) (0.815)	0.861 (0.191) (0.211) (0.186)	0.885 (0.319) (0.339) (0.286)	0.890 (0.297) (0.318) (0.275)	97	97	97	95
5	1.013 (0.901) (0.906) (0.887)	0.843 (0.104) (0.125) (0.090)	0.835 (0.129) (0.151) (0.098)	0.842 (0.144) (0.168) (0.108)	97	97	97	94

History: Five Qtrs After Initial Release
1985:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.771 (0.001) (0.002) (0.005)	0.700 (0.001) (0.001) (0.011)	0.700 (0.001) (0.001) (0.011)	0.759 (0.000) (0.000) (0.001)	145	145	145	142
2	0.879 (0.025) (0.028) (0.067)	0.762 (0.001) (0.001) (0.003)	0.817 (0.003) (0.004) (0.009)	0.782 (0.001) (0.001) (0.002)	145	145	145	142
3	0.911 (0.255) (0.265) (0.255)	0.784 (0.006) (0.008) (0.010)	0.815 (0.038) (0.043) (0.035)	0.827 (0.042) (0.048) (0.038)	145	145	145	142
4	0.962 (0.660) (0.668) (0.633)	0.824 (0.042) (0.049) (0.039)	0.844 (0.099) (0.109) (0.077)	0.856 (0.086) (0.096) (0.071)	145	145	145	142
5	0.969 (0.707) (0.717) (0.671)	0.822 (0.026) (0.033) (0.021)	0.807 (0.031) (0.039) (0.019)	0.821 (0.043) (0.052) (0.026)	145	145	145	142

History: Five Qtrs After Initial Release
1985:01-1996:04

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.712 (0.000) (0.001) (0.000)	0.667 (0.001) (0.002) (0.000)	0.667 (0.001) (0.002) (0.000)	0.680 (0.013) (0.018) (0.003)	48	48	48	47
2	0.729 (0.000) (0.000) (0.000)	0.654 (0.005) (0.009) (0.004)	0.672 (0.010) (0.016) (0.010)	0.615 (0.011) (0.018) (0.000)	48	48	48	47
3	0.748 (0.000) (0.000) (0.000)	0.663 (0.023) (0.036) (0.019)	0.680 (0.046) (0.065) (0.035)	0.696 (0.047) (0.067) (0.018)	48	48	48	47
4	0.814 (0.001) (0.004) (0.004)	0.724 (0.033) (0.054) (0.027)	0.735 (0.070) (0.100) (0.050)	0.762 (0.029) (0.049) (0.015)	48	48	48	47
5	0.855 (0.157) (0.206) (0.126)	0.763 (0.053) (0.086) (0.043)	0.724 (0.047) (0.079) (0.032)	0.759 (0.062) (0.097) (0.037)	48	48	48	48

History: Five Qtrs After Initial Release
1997:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.800 (0.043) (0.047) (0.091)	0.715 (0.019) (0.022) (0.090)	0.715 (0.019) (0.022) (0.090)	0.800 (0.009) (0.010) (0.031)	97	97	97	95
2	0.938 (0.377) (0.386) (0.475)	0.801 (0.021) (0.025) (0.044)	0.874 (0.073) (0.081) (0.119)	0.854 (0.023) (0.027) (0.075)	97	97	97	95
3	0.968 (0.753) (0.760) (0.753)	0.824 (0.057) (0.067) (0.077)	0.860 (0.187) (0.201) (0.184)	0.869 (0.196) (0.211) (0.199)	97	97	97	95
4	1.014 (0.902) (0.906) (0.893)	0.856 (0.176) (0.195) (0.170)	0.880 (0.298) (0.318) (0.265)	0.885 (0.273) (0.294) (0.251)	97	97	97	95
5	1.006 (0.953) (0.955) (0.946)	0.839 (0.096) (0.116) (0.083)	0.833 (0.125) (0.147) (0.094)	0.840 (0.138) (0.161) (0.102)	97	97	97	94

History: Nine Qtrs After Initial Release
1985:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.760 (0.001) (0.001) (0.004)	0.687 (0.001) (0.001) (0.009)	0.687 (0.001) (0.001) (0.009)	0.745 (0.000) (0.000) (0.001)	145	145	145	142
2	0.876 (0.024) (0.027) (0.066)	0.758 (0.001) (0.001) (0.003)	0.815 (0.004) (0.005) (0.010)	0.777 (0.001) (0.001) (0.002)	145	145	145	142
3	0.906 (0.230) (0.240) (0.231)	0.779 (0.006) (0.007) (0.009)	0.811 (0.035) (0.041) (0.033)	0.825 (0.042) (0.048) (0.038)	145	145	145	142
4	0.956 (0.604) (0.614) (0.575)	0.819 (0.037) (0.044) (0.034)	0.839 (0.090) (0.101) (0.069)	0.851 (0.077) (0.087) (0.063)	145	145	145	142
5	0.970 (0.716) (0.725) (0.681)	0.821 (0.027) (0.033) (0.021)	0.804 (0.030) (0.038) (0.018)	0.820 (0.043) (0.052) (0.026)	145	145	145	142

History: Nine Qtrs After Initial Release
1985:01-1996:04

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.703 (0.000) (0.000) (0.000)	0.647 (0.001) (0.003) (0.000)	0.647 (0.001) (0.003) (0.000)	0.658 (0.010) (0.015) (0.002)	48	48	48	47
2	0.723 (0.000) (0.000) (0.000)	0.646 (0.006) (0.010) (0.005)	0.664 (0.013) (0.021) (0.014)	0.601 (0.012) (0.019) (0.001)	48	48	48	47
3	0.741 (0.000) (0.000) (0.000)	0.658 (0.022) (0.035) (0.018)	0.675 (0.044) (0.063) (0.034)	0.698 (0.053) (0.074) (0.021)	48	48	48	47
4	0.804 (0.001) (0.003) (0.003)	0.717 (0.033) (0.054) (0.026)	0.728 (0.069) (0.098) (0.048)	0.756 (0.028) (0.048) (0.015)	48	48	48	47
5	0.861 (0.191) (0.242) (0.155)	0.764 (0.061) (0.097) (0.048)	0.723 (0.050) (0.083) (0.034)	0.759 (0.068) (0.105) (0.042)	48	48	48	48

History: Nine Qtrs After Initial Release
1997:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.788 (0.033) (0.037) (0.076)	0.705 (0.018) (0.020) (0.084)	0.705 (0.018) (0.020) (0.084)	0.791 (0.006) (0.008) (0.025)	97	97	97	95
2	0.934 (0.354) (0.364) (0.454)	0.797 (0.018) (0.022) (0.041)	0.872 (0.073) (0.081) (0.120)	0.852 (0.021) (0.026) (0.075)	97	97	97	95
3	0.964 (0.724) (0.731) (0.724)	0.819 (0.053) (0.062) (0.073)	0.857 (0.180) (0.195) (0.178)	0.865 (0.187) (0.202) (0.191)	97	97	97	95
4	1.009 (0.936) (0.938) (0.929)	0.852 (0.164) (0.183) (0.160)	0.876 (0.286) (0.306) (0.253)	0.881 (0.260) (0.280) (0.238)	97	97	97	95
5	1.005 (0.964) (0.965) (0.958)	0.838 (0.095) (0.114) (0.081)	0.830 (0.122) (0.143) (0.091)	0.838 (0.135) (0.158) (0.099)	97	97	97	94

History: Latest Vintage
1985:01-2021:01

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.759 (0.001) (0.001) (0.005)	0.683 (0.001) (0.001) (0.010)	0.683 (0.001) (0.001) (0.010)	0.743 (0.000) (0.000) (0.001)	145	145	145	142
2	0.876 (0.023) (0.026) (0.066)	0.759 (0.001) (0.001) (0.003)	0.815 (0.004) (0.005) (0.010)	0.777 (0.001) (0.001) (0.002)	145	145	145	142
3	0.905 (0.222) (0.232) (0.223)	0.780 (0.005) (0.007) (0.009)	0.811 (0.036) (0.041) (0.034)	0.825 (0.042) (0.048) (0.039)	145	145	145	142
4	0.952 (0.569) (0.579) (0.538)	0.817 (0.034) (0.041) (0.032)	0.837 (0.085) (0.095) (0.065)	0.852 (0.079) (0.089) (0.065)	145	145	145	142
5	0.969 (0.712) (0.721) (0.676)	0.821 (0.027) (0.033) (0.021)	0.804 (0.030) (0.037) (0.018)	0.819 (0.041) (0.050) (0.024)	145	145	145	142

History: Latest Vintage
1985:01-1996:04

H	RMSE (S/NC)	RMSE (S/IAR)	RMSE (S/DAR)	RMSE (S/DARM)	N1	N2	N3	N4
1	0.702 (0.000) (0.000) (0.000)	0.640 (0.001) (0.002) (0.000)	0.640 (0.001) (0.002) (0.000)	0.655 (0.012) (0.016) (0.002)	48	48	48	47
2	0.717 (0.000) (0.000) (0.000)	0.642 (0.005) (0.009) (0.004)	0.660 (0.012) (0.019) (0.013)	0.598 (0.013) (0.020) (0.000)	48	48	48	47
3	0.736 (0.000) (0.000) (0.000)	0.657 (0.021) (0.034) (0.019)	0.674 (0.045) (0.063) (0.035)	0.697 (0.052) (0.072) (0.022)	48	48	48	47
4	0.793 (0.000) (0.002) (0.002)	0.713 (0.030) (0.050) (0.024)	0.723 (0.063) (0.091) (0.044)	0.760 (0.030) (0.051) (0.016)	48	48	48	47
5	0.859 (0.174) (0.224) (0.140)	0.764 (0.066) (0.102) (0.051)	0.723 (0.052) (0.084) (0.035)	0.756 (0.059) (0.093) (0.035)	48	48	48	48

History: Latest Vintage
1997:01-2021:01

H	RMSE(S/NC)	RMSE(S/IAR)	RMSE(S/DAR)	RMSE(S/DARM)	N1	N2	N3	N4
1	0.785 (0.035) (0.039) (0.080)	0.702 (0.019) (0.021) (0.088)	0.702 (0.019) (0.021) (0.088)	0.788 (0.007) (0.009) (0.029)	97	97	97	95
2	0.936 (0.367) (0.377) (0.469)	0.799 (0.019) (0.023) (0.043)	0.874 (0.076) (0.083) (0.125)	0.853 (0.021) (0.025) (0.076)	97	97	97	95
3	0.965 (0.734) (0.741) (0.733)	0.820 (0.053) (0.062) (0.073)	0.858 (0.183) (0.197) (0.180)	0.867 (0.191) (0.206) (0.195)	97	97	97	95
4	1.009 (0.936) (0.938) (0.929)	0.851 (0.162) (0.180) (0.157)	0.876 (0.284) (0.304) (0.251)	0.881 (0.260) (0.280) (0.238)	97	97	97	95
5	1.005 (0.961) (0.963) (0.956)	0.838 (0.094) (0.114) (0.081)	0.830 (0.120) (0.142) (0.090)	0.838 (0.134) (0.157) (0.099)	97	97	97	94

Notes for Table 1B.

- (1) The forecast horizon is given by H, where H = 1 is the SPF forecast for the current quarter.
- (2) The headers RMSE(S/NC), RMSE(S/IAR), RMSE(S/DAR), and RMSE(S/DARM) are the ratios of the SPF root-mean-square error to that of the benchmark models: No-change (NC), indirect autoregression (IAR), direct autoregression (DAR), and direct autoregression augmented with monthly information (DARM). These statistics may differ slightly from those reported in Table 1A because they incorporate only those observations common to both the SPF and the benchmark model. The previous statistics make use of all available observations for each model.
- (3) All models are estimated on a rolling window of 60 observations from the Philadelphia Fed real-time data set.
- (4) A set of three two-sided p-values (in parentheses) accompanies each statistic. These are the p-values for the test of the equality of mean-square-error. The first is for the Diebold-Mariano (1995, JBES) statistic, using a uniform lag window with the truncation lag set to the forecast horizon minus one. (The tables report the p-values using a Bartlett window when the uniform window produces a negative standard error.) The second is for the Harvey-Leybourne-Newbold (1997, IJF) correction to the Diebold-Mariano statistic. The third is for the Diebold-Mariano statistic, using a Bartlett lag window with the truncation lag increased four quarters.
- (5) The headers N1, N2, N3, and N4 show the number of observations used in constructing each ratio of root-mean-square errors.
- (6) 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)
By Forecast Step (1 to 5)
Transformation: Level

Last Updated: 08/21/2023 13:25

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2016:04	1.171	1.220	1.234	1.280	1.300
2017:01	1.219	1.201	1.240	1.261	1.300
2017:02	1.268	1.250	1.237	1.272	1.315
2017:03	1.225	1.290	1.269	1.246	1.286
2017:04	1.207	1.240	1.295	1.284	1.256
2018:01	1.269	1.246	1.263	1.320	1.300
2018:02	1.319	1.275	1.258	1.280	1.340
2018:03	1.284	1.319	1.278	1.270	1.300
2018:04	1.256	1.298	1.335	1.290	1.285
2019:01	1.234	1.274	1.306	1.350	1.294
2019:02	1.208	1.248	1.282	1.324	1.365
2019:03	1.267	1.226	1.266	1.293	1.334
2019:04	1.290	1.269	1.240	1.270	1.303
2020:01	1.369	1.281	1.266	1.255	1.270
2020:02	0.950	1.376	1.284	1.278	1.260
2020:03	1.196	0.994	1.377	1.284	1.278
2020:04	1.418	1.216	1.107	1.374	1.285
2021:01	1.550	1.410	1.260	1.208	1.393
2021:02	1.630	1.514	1.434	1.288	1.241
2021:03	1.595	1.624	1.518	1.450	1.304
2021:04	1.589	1.600	1.587	1.510	1.437
2022:01	1.650	1.599	1.595	1.563	1.530
2022:02	1.720	1.655	1.601	1.587	1.612
2022:03	1.590	1.668	1.649	1.594	1.593
2022:04	1.430	1.550	1.650	1.650	1.583
2023:01	1.377	1.400	1.526	1.635	1.646
2023:02	1.360	1.350	1.380	1.545	1.601
2023:03	1.400	1.325	1.333	1.400	1.528
2023:04	NA	1.400	1.316	1.347	1.400
2024:01	NA	NA	1.393	1.332	1.368
2024:02	NA	NA	NA	1.395	1.378
2024:03	NA	NA	NA	NA	1.414

Notes for Table 2.

- (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)
 By Forecast Step (1 to 5)
 Transformation: Level
 Lag Length for IAR(p): AIC

Source for Historical Realizations: Census Bureau via Haver Analytics

Last Updated: 08/21/2023 13:25

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2016:04	1.135	1.158	1.178	1.199	1.286
2017:01	1.230	1.138	1.163	1.186	1.211
2017:02	1.263	1.239	1.136	1.165	1.198
2017:03	1.166	1.278	1.252	1.132	1.166
2017:04	1.172	1.173	1.298	1.272	1.131
2018:01	1.245	1.153	1.155	1.301	1.280
2018:02	1.317	1.232	1.135	1.138	1.303
2018:03	1.265	1.333	1.241	1.130	1.134
2018:04	1.228	1.287	1.358	1.253	1.123
2019:01	1.228	1.224	1.289	1.370	1.247
2019:02	1.165	1.197	1.196	1.271	1.369
2019:03	1.253	1.141	1.176	1.175	1.266
2019:04	1.274	1.239	1.119	1.161	1.165
2020:01	1.473	1.274	1.239	1.106	1.142
2020:02	1.476	1.494	1.273	1.238	1.088
2020:03	0.910	1.478	1.508	1.266	1.224
2020:04	1.350	0.873	1.494	1.526	1.252
2021:01	1.365	1.195	0.864	1.477	1.519
2021:02	1.546	1.310	1.313	0.865	1.451
2021:03	1.522	1.487	1.471	1.233	0.867
2021:04	1.530	1.480	1.434	1.400	1.167
2022:01	1.618	1.497	1.441	1.386	1.256
2022:02	1.741	1.594	1.466	1.406	1.343
2022:03	1.636	1.729	1.570	1.437	1.373
2022:04	1.445	1.619	1.718	1.548	1.410
2023:01	1.398	1.436	1.603	1.706	1.527
2023:02	1.393	1.391	1.424	1.588	1.695
2023:03	1.443	1.388	1.383	1.413	1.573
2023:04	NA	1.436	1.384	1.375	1.402
2024:01	NA	NA	1.434	1.379	1.367
2024:02	NA	NA	NA	1.432	1.375
2024:03	NA	NA	NA	NA	1.430

Notes for Table 3.

- (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)
By Forecast Step (1 to 5)
Transformation: Level

Source for Historical Realizations: Census Bureau via Haver Analytics

Last Updated: 08/21/2023 13:25

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2016:04	1.138	1.160	1.133	1.133	1.163
2017:01	1.216	1.138	1.160	1.133	1.133
2017:02	1.253	1.216	1.138	1.160	1.133
2017:03	1.164	1.253	1.216	1.138	1.160
2017:04	1.165	1.164	1.253	1.216	1.138
2018:01	1.251	1.165	1.164	1.253	1.216
2018:02	1.318	1.251	1.165	1.164	1.253
2018:03	1.262	1.318	1.251	1.165	1.164
2018:04	1.218	1.262	1.318	1.251	1.165
2019:01	1.236	1.218	1.262	1.318	1.251
2019:02	1.193	1.236	1.218	1.262	1.318
2019:03	1.263	1.193	1.236	1.218	1.262
2019:04	1.282	1.263	1.193	1.236	1.218
2020:01	1.441	1.282	1.263	1.193	1.236
2020:02	1.466	1.441	1.282	1.263	1.193
2020:03	1.044	1.466	1.441	1.282	1.263
2020:04	1.430	1.044	1.466	1.441	1.282
2021:01	1.592	1.430	1.044	1.466	1.441
2021:02	1.613	1.592	1.430	1.044	1.466
2021:03	1.568	1.613	1.592	1.430	1.044
2021:04	1.566	1.568	1.613	1.592	1.430
2022:01	1.644	1.566	1.568	1.613	1.592
2022:02	1.753	1.644	1.566	1.568	1.613
2022:03	1.652	1.753	1.644	1.566	1.568
2022:04	1.461	1.652	1.753	1.644	1.566
2023:01	1.403	1.461	1.652	1.753	1.644
2023:02	1.395	1.403	1.461	1.652	1.753
2023:03	1.447	1.395	1.403	1.461	1.652
2023:04	NA	1.447	1.395	1.403	1.461
2024:01	NA	NA	1.447	1.395	1.403
2024:02	NA	NA	NA	1.447	1.395
2024:03	NA	NA	NA	NA	1.447

Notes for Table 4.

- (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)
By Forecast Step (1 to 5)
Transformation: Level
Lag Length for DAR(p): AIC

Source for Historical Realizations: Census Bureau via Haver Analytics

Last Updated: 08/21/2023 13:25

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2016:04	1.135	1.157	1.179	1.195	1.296
2017:01	1.230	1.142	1.229	1.267	1.277
2017:02	1.263	1.236	1.141	1.178	1.194
2017:03	1.166	1.280	1.250	1.137	1.278
2017:04	1.172	1.179	1.300	1.272	1.160
2018:01	1.245	1.129	1.170	1.299	1.266
2018:02	1.317	1.254	1.161	1.174	1.328
2018:03	1.265	1.318	1.275	1.183	1.203
2018:04	1.228	1.292	1.322	1.257	1.153
2019:01	1.228	1.234	1.262	1.328	1.278
2019:02	1.165	1.231	1.225	1.296	1.346
2019:03	1.253	1.157	1.232	1.237	1.283
2019:04	1.274	1.211	1.136	1.221	1.227
2020:01	1.473	1.254	1.189	1.110	1.234
2020:02	1.476	1.454	1.239	1.188	1.128
2020:03	0.910	1.486	1.436	1.236	1.171
2020:04	1.350	1.049	1.502	1.443	1.236
2021:01	1.365	1.346	0.989	1.471	1.422
2021:02	1.546	1.419	1.428	0.960	1.489
2021:03	1.522	1.524	1.625	1.446	1.053
2021:04	1.530	1.405	1.470	1.442	1.309
2022:01	1.618	1.747	1.446	1.451	1.526
2022:02	1.741	1.633	1.738	1.487	1.510
2022:03	1.636	1.689	1.635	1.679	1.668
2022:04	1.445	1.623	1.691	1.547	1.440
2023:01	1.398	1.418	1.576	1.667	1.430
2023:02	1.393	1.347	1.345	1.636	1.617
2023:03	1.443	1.371	1.280	1.314	1.433
2023:04	NA	1.428	1.344	1.340	1.222
2024:01	NA	NA	1.418	1.351	1.302
2024:02	NA	NA	NA	1.423	1.332
2024:03	NA	NA	NA	NA	1.408

Notes for Table 5.

- (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)
By Forecast Step (1 to 5)
Transformation: Level
Lag Length for DARM(p): AIC

Source for Historical Realizations: Census Bureau via Haver Analytics

Last Updated: 08/21/2023 13:25

Qtr Forecast	Step 1	Step 2	Step 3	Step 4	Step 5
2016:04	1.093	1.177	1.197	1.209	1.293
2017:01	1.244	1.077	1.191	1.243	1.266
2017:02	1.240	1.241	1.086	1.202	1.212
2017:03	1.194	1.250	1.296	1.064	1.225
2017:04	1.165	1.211	1.255	1.291	1.040
2018:01	1.212	1.145	1.203	1.252	1.310
2018:02	1.310	1.187	1.154	1.224	1.258
2018:03	1.230	1.312	1.184	1.172	1.265
2018:04	1.228	1.253	1.320	1.179	1.138
2019:01	NA	1.256	1.217	1.316	1.191
2019:02	1.157	NA	1.214	1.216	1.328
2019:03	1.249	1.110	NA	1.237	1.195
2019:04	1.257	1.214	1.123	NA	1.218
2020:01	1.517	1.261	1.191	1.061	NA
2020:02	1.435	1.515	1.215	1.200	1.106
2020:03	1.119	1.390	1.463	1.226	1.168
2020:04	1.394	1.011	1.454	1.479	1.186
2021:01	1.449	1.347	1.043	1.489	1.506
2021:02	1.600	1.445	1.423	0.962	1.418
2021:03	1.719	1.544	1.643	1.447	1.069
2021:04	1.515	1.484	1.549	1.428	1.309
2022:01	1.658	1.726	1.564	1.471	1.525
2022:02	1.759	1.665	1.705	1.608	1.484
2022:03	1.548	1.707	1.638	1.645	1.712
2022:04	1.470	1.527	1.693	1.588	1.449
2023:01	1.406	1.447	1.609	1.689	1.542
2023:02	1.403	1.360	1.449	1.560	1.696
2023:03	1.441	1.403	1.274	1.324	1.468
2023:04	NA	1.455	1.369	1.312	1.273
2024:01	NA	NA	1.406	1.380	1.268
2024:02	NA	NA	NA	1.406	1.375
2024:03	NA	NA	NA	NA	1.442

Notes for Table 6.

- (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)
Philadelphia Fed Real-Time Data Set

Variable: HOUSING (Housing Starts)
Transformation: Level

Source for Historical Realizations: Census Bureau via Haver Analytics

Last Updated: 08/21/2023 13:25

Column (1): Initial Release
Column (2): One Qtr After Initial Release
Column (3): Five Qtrs After Initial Release
Column (4): Nine Qtrs After Initial Release
Column (5): Latest Vintage

Obs. Date	(1)	(2)	(3)	(4)	(5)
2016:04	1.216	1.248	1.248	1.253	1.235
2017:01	1.253	1.238	1.231	1.231	1.217
2017:02	1.164	1.167	1.171	1.169	1.184
2017:03	1.165	1.172	1.172	1.175	1.182
2017:04	1.251	1.256	1.259	1.260	1.236
2018:01	1.318	1.317	1.321	1.312	1.298
2018:02	1.262	1.261	1.260	1.264	1.275
2018:03	1.218	1.234	1.233	1.240	1.249
2018:04	1.236	1.185	1.185	1.176	1.166
2019:01	1.193	1.213	1.204	1.196	1.185
2019:02	1.263	1.256	1.257	1.274	1.267
2019:03	1.282	1.282	1.288	1.293	1.304
2019:04	1.441	1.441	1.433	1.405	1.410
2020:01	1.466	1.484	1.485	1.470	1.468
2020:02	1.044	1.079	1.086	1.087	1.082
2020:03	1.430	1.432	1.440	1.449	1.456
2020:04	1.592	1.584	1.575	1.574	1.582
2021:01	1.613	1.599	1.581	1.575	1.575
2021:02	1.568	1.588	1.591	NA	1.582
2021:03	1.566	1.562	1.569	NA	1.576
2021:04	1.644	1.670	1.679	NA	1.690
2022:01	1.753	1.720	1.718	NA	1.718
2022:02	1.652	1.647	NA	NA	1.636
2022:03	1.461	1.450	NA	NA	1.446
2022:04	1.403	1.398	NA	NA	1.405
2023:01	1.395	1.385	NA	NA	1.385
2023:02	1.447	NA	NA	NA	1.447
2023:03	NA	NA	NA	NA	NA
2023:04	NA	NA	NA	NA	NA
2024:01	NA	NA	NA	NA	NA
2024:02	NA	NA	NA	NA	NA
2024:03	NA	NA	NA	NA	NA

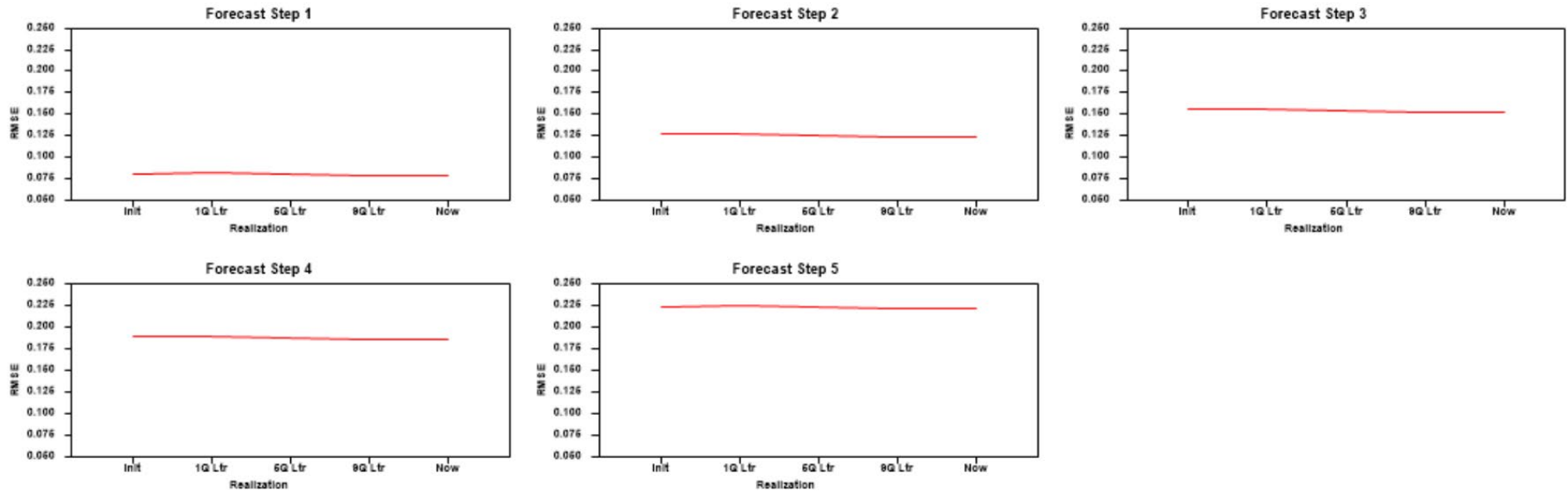
Notes for Table 7.

- (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: 1985:01-2021:01

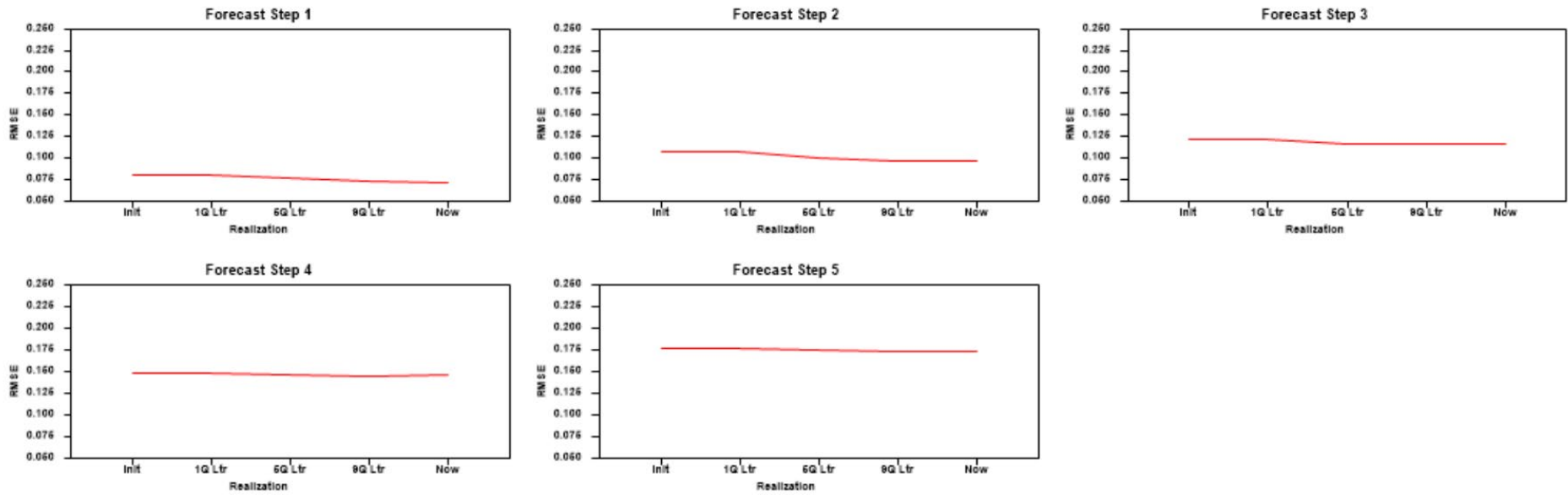
SPF Projections for Housing Starts, Transformation: Level



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

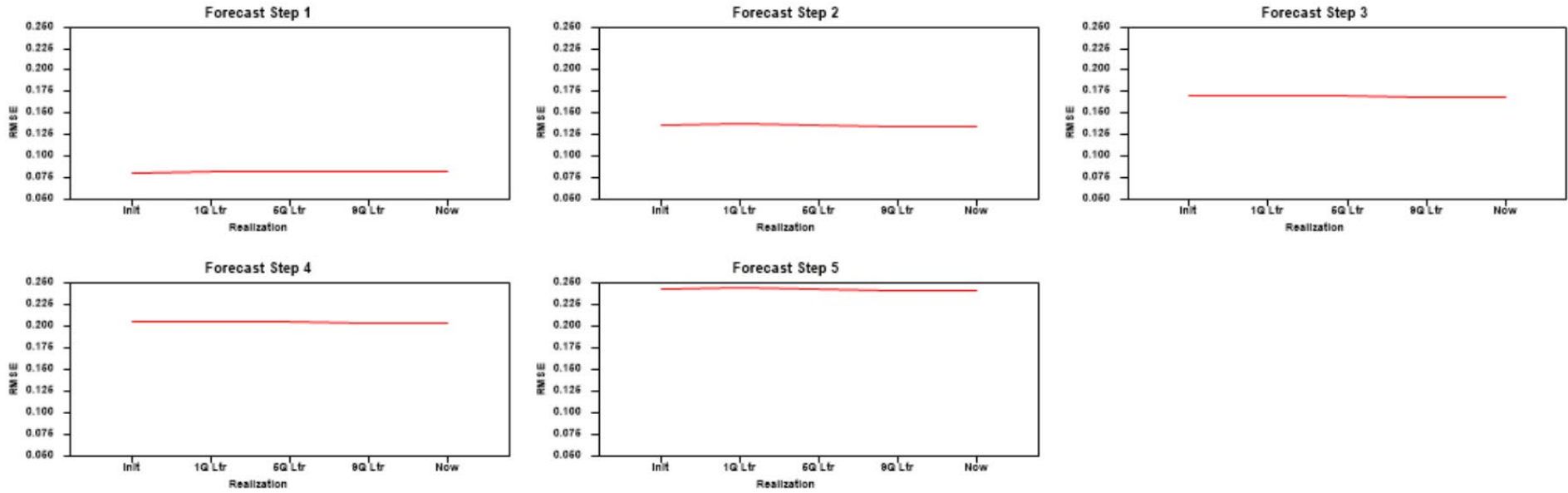
SPF Projections for Housing Starts, Transformation: Level



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: 1997:01-2021:01

SPF Projections for Housing Starts, Transformation: Level



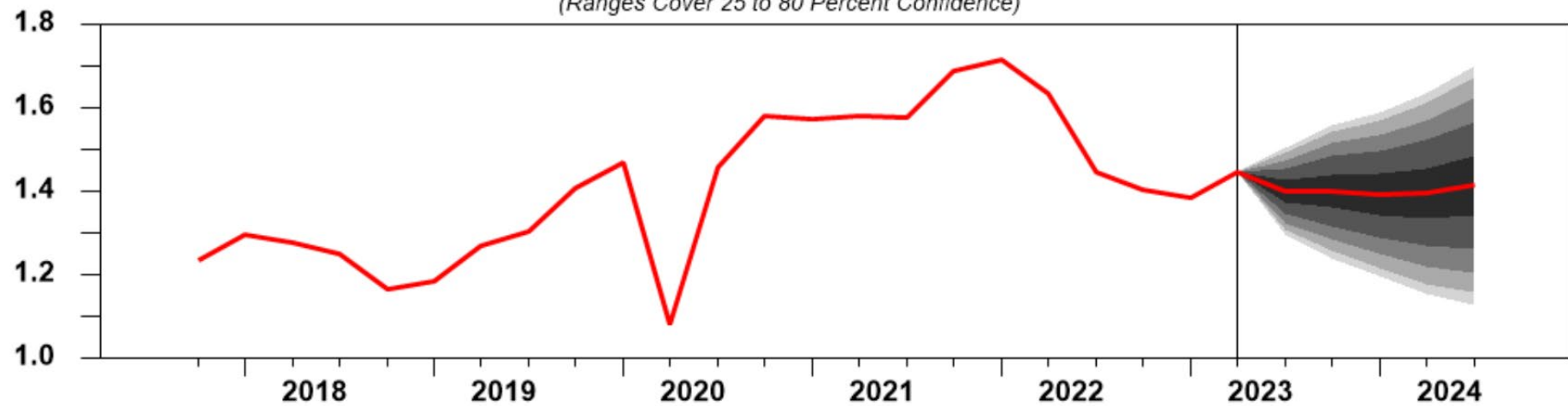
The RMSE is plotted against the realization used to compute it, from the value on Initial release to the value as we now know it. Source: Tom Stark, FRB Philadelphia.

Housing Starts

History, Forecasts, and Ranges for the SPF of 2023:03

Level

(Ranges Cover 25 to 80 Percent Confidence)



Ranges at each horizon use the $N(0, \text{MSE})$ density. The MSEs are based on the sample 85:01-22:01 and use the realization: Five Qtrs After Initial Release. Source: T.Stark, FRB Phila.