**Finance and Economics Discussion Series**
**Divisions of Research & Statistics and Monetary Affairs**
**Federal Reserve Board, Washington, D.C.**

# The Accuracy of Forecasts Prepared for the Federal Open Market Committee

**Andrew C. Chang and Tyler J. Hanson**

**2015-062**

Please cite this paper as:
Chang, Andrew C. and Tyler J. Hanson (2015). "The Accuracy of Forecasts Prepared for the Federal Open Market Committee," Finance and Economics Discussion Series 2015-062. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.062.

# The Accuracy of Forecasts Prepared for the Federal Open Market Committee

Andrew C. Chang* and Tyler J. Hanson†

July 9, 2015

**Abstract**

We analyze forecasts of consumption, nonresidential investment, residential investment, government spending, exports, imports, inventories, gross domestic product, inflation, and unemployment prepared by the staff of the Board of Governors of the Federal Reserve System for meetings of the Federal Open Market Committee from 1997 to 2008, called the Greenbooks. We compare the root mean squared error, mean absolute error, and the proportion of directional errors of Greenbook forecasts of these macroeconomic indicators to the errors from three forecasting benchmarks: a random walk, a first-order autoregressive model, and a Bayesian model averaged forecast from a suite of univariate time-series models commonly taught to first-year economics graduate students. We estimate our forecasting benchmarks both on end-of-sample vintage and real-time vintage data. We find find that Greenbook forecasts significantly outperform our benchmark forecasts for horizons less than one quarter ahead. However, by the one-year forecast horizon, typically at least one of our forecasting benchmarks performs as well as Greenbook forecasts. Greenbook forecasts of the personal consumption expenditures and unemployment tend to do relatively well, while Greenbook forecasts of inventory investment, government expenditures, and inflation tend to do poorly.

JEL Codes: C53; E17; E27; E37; F17

Keywords: Bayesian Model Averaging; Federal Open Market Committee; Forecast Accuracy; Greenbook; National Income and Product Accounts; NIPA; Real-Time Data

*Chang: Board of Governors of the Federal Reserve System. 20th St. NW and Constitution Ave., Washington DC 20551 USA. +1 (657) 464-3286. a.christopher.chang@gmail.com. https://sites.google.com/site/andrewchristopherchang/.

†Hanson: Hopper. 275 Third Street, Cambridge, MA 02142 USA. thanson2691@gmail.com. The views and opinions expressed here are those of the authors and are not necessarily those of the Board of Governors of the Federal Reserve System or Hopper. We thank Stephanie Aaronson, David Lebow, Paul Lengermann, Phillip Li, Priyanka Shahane, and Missaka Warusawitharana for helpful comments. We thank Kim T. Mai for research assistance. Any errors are ours.

# 1 Introduction

Accurate assessments of the real-time state of economic activity and accurate forecasts of the future path of activity are important inputs for monetary policy decisions. Central banks invest considerable resources in forecasting economic activity to guide policy decisions. For example, prior to meetings of the Federal Open Market Committee (FOMC), the Federal Reserve Board staff prepare a detailed projection of US economic activity for the FOMC, known as the Greenbook.[1] Production of the Greenbook employs around a hundred economists and research assistants in addition to other editorial, legal, and administrative staff.[2] Despite the considerable effort that goes into Greenbook production because of the Greenbook's contribution to monetary policy decisions, significant uncertainty surrounds Greenbook forecasts (Reifschneider and Tulip, 2007; Tulip, 2009).

Our primary contribution is analyzing the accuracy of Greenbook forecasts of 10 key aggregates of the US economy in a unified framework, as opposed to only gross domestic product (GDP) or inflation (Romer and Romer, 2000; Faust and Wright, 2009; Wright, 2009; Tulip, 2009; Arai, 2014). In addition to these two key macroeconomic indicators, we analyze the unemployment rate and the major components of GDP from the National Income and Product Accounts (NIPA): consumption, nonresidential investment, residential investment, government spending, exports, imports, and business inventories. We consider forecasts from 1997 to 2008.

We compare the accuracy of Greenbook forecasts to the accuracy of forecasts from three benchmark reduced-form univariate methods: a random walk, a first-order autoregressive (AR) model, and a Bayesian model averaged forecast from a pool of univariate time-series models taught in first-year economics graduate courses. We choose these benchmarks because of their parsimony, ease of implementation, and independence from auxiliary data. We assess whether the Greenbook forecasts, which require substantially more resources to prepare than any of these methods, empirically outperform these simple forecasts. Our dependence only on simple univariate methods also allows us to use only models that were available to forecasters at the time the forecasts were

[1] Since 2010 this projection is called the Tealbook.

[2] As of this writing, there are approximately forty economists and research assistants are formally assigned to Greenbook preparation, but many more participants are informally involved.

generated, which reduces potential hindsight bias in model selection (Tulip, 2009). We measure accuracy as root mean squared error (RMSE), mean absolute error (MAE), and the proportion of forecasts where the predicted sign of the acceleration is incorrect, which we call mean directional error (MDE).

To avoid the pitfalls of conducting pseudo out-of-sample forecasting exercises on current-vintage data, we estimate our three benchmarks using two classes of data available to Greenbook forecasters at the time the forecasts were generated.[3] For the first class of data, we estimate models using the "conventional" data that professional forecasters employ, or what Koenig, Dolmas, and Piger (2003) refer to as *end-of-sample vintage* (EOS) data. These data are the fully revised version of a series at a given point in time. For example, to forecast GDP growth for 2000 Q1, we estimate models using the latest-revised data available as of 1999 Q4. To forecast GDP growth for 2000 Q2, we estimate models using the latest-revised data available as of 2000 Q1, and so on. Because of the practice of US statistical agencies of continually revising previously published estimates, the older datapoints in EOS data have undergone more revisions than more recent datapoints.

For the second class of data, we estimate models on *real-time vintage* (RTV) data, a time series of datapoints where each datapoint has undergone the same number of data revisions. For example, to estimate the third-release (twice-revised) estimate of GDP growth for 2000 Q1 using a univariate model in GDP with RTV data, the right-hand side observations consist of only previous third-release (twice-revised) estimates of GDP growth. In contrast with EOS data, older RTV datapoints have undergone the identical number of data revisions as the newer datapoints.[4]

We find that Greenbook forecasts significantly outperform our benchmark forecasts in the very near term, typically for forecast horizons within one quarter. This performance carries through whether we measure performance by RMSE, MAE, or MDE. However, by the one-year forecast horizon, typically at least one of our forecasting benchmarks performs as well as Greenbook fore-

[3]Estimating models using *current-vintage* data, the fully revised versions of data that are available today, can skew the forecasting performance of models with information not available to forecasters at the time forecasts were actually generated (Koenig, Dolmas, and Piger, 2003; Reifschneider and Tulip, 2007; Tulip, 2009; Clements and Galvão, 2013).

[4]The third-release (twice-revised) estimate is also called the "final" estimate.

casts.

There is some sector heterogeneity of forecast performance. The Greenbook forecasts of the unemployment rate and personal consumption expenditures (PCE) tend to outperform our benchmarks for longer forecast horizons. The Greenbook forecasts of the change in business inventories, core PCE inflation, and government spending tend to perform similarly to or are outperformed by our benchmarks at shorter forecast horizons.

## 2 Data and Sample Frame

We obtain historical unemployment rates and NIPA data from the St. Louis Federal Reserve's archival database (ALFRED). In addition to analyzing total GDP, we also consider PCE, nonresidential private fixed investment (NRPFI), residential private fixed investment (RES), government expenditures (GOV), change in business inventories (CBI), exports, imports, and core PCE inflation. Most NIPA series are quarterly and the Greenbook contains forecasts on a quarterly basis. For core PCE inflation and unemployment, where data are available monthly, we convert monthly variables to quarterly variables by averaging monthly values.

We use Greenbook forecasts from 1997 to 2008 (1997 is the earliest full year in which all series are available, and 2008 is the latest year of Greenbook forecasts that have been made public as of this writing). Greenbook forecasts are available from the publicly available Domestic Economic Developments and Outlook texts on the Federal Reserve Board's website (Federal Reserve Board, 2014).

Because Greenbooks are published at irregular intervals that do not correspond directly to calendar months or quarters, we use the final Greenbook released in each quarter. We estimate our benchmark forecasts using the vintage data series as of the Greenbook's day of release, giving equal information sets to both our benchmark forecasts and the Greenbook releases. Our estimation sample period begins in the first quarter of 1986, a year that succeeds most estimates for the beginning of the "great moderation" and falls just after a NIPA benchmark revision at the end of

1985. We use real seasonally adjusted annual percent changes for all series except for business inventories, where we use annualized real seasonally adjusted first-differences, and the unemployment rate, where we use levels. We choose these units to match Greenbook forecasts.

To compute loss functions, for GDP and its components we take then Bureau of Economic Analysis's (BEA) third-release (twice-revised) estimate as the forecasting target. For the loss functions for core PCE inflation and the unemployment rate, we use the quarterly average computed when the last month of data for the quarter is first available.[5] The results are largely robust to using the December 2014 vintage of data as the forecasting target. Table 1 shows summary statistics of our data.

# 3 Comparison Models

To establish a relevant comparison between Greenbook forecasts and univariate forecasts, we first estimate three naïve benchmark forecasts: the historical mean, a random walk, and a AR(1) model. Let $y$ be our variable of interest and $\hat{y}$ be a forecast of $y$ from a particular model. The random walk is specified as:

$$y_t = y_{t-1} + \varepsilon_t \tag{1}$$

such that the forecast for each horizon is equal to the last observed value, which we take as the observed value two quarters prior to the Greenbook publication quarter.[6] We specify the AR(1) model with a constant:

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t \tag{2}$$

For our final benchmark, we construct a Bayesian model averaged (BMA) forecast. We use

[5]Monthly estimates of core PCE inflation and the unemployment rate are available in the subsequent month. Therefore, this procedure yields quarterly averages computed with the core PCE inflation and unemployment releases in January, April, July, and October.

[6]Two quarters prior to publication is the first quarter where the BEA has published its third-release (twice-revised) estimate for the quarter for every Greenbook in the sample.

Bayesian weights over 43 univariate forecasting models, using the Schwarz Bayesian Information Criterion (SBIC) approximation for the log marginal likelihood proposed by Raftery (1995).

The BMA forecast $\bar{y}$ is represented as:

$$\bar{y}_t = \sum_{i=1}^{N} \hat{y}_{i,t} Pr(M_i \mid y) \tag{3}$$

where $Pr(M_i \mid y)$ represents the probability that model $i$ is true given the data. This probability can be approximated using the SBIC such that:

$$Pr(M_i \mid y) = \frac{e^{SBIC_i} Pr(M_i)}{\sum_{i=1}^{N} e^{SBIC_i} Pr(M_i)} \tag{4}$$

where $Pr(M_i)$ represents the prior model probability for model $i$. For our weighted forecast, we assign equal prior probability to all 43 model specifications, so $Pr(M_i) = 1/43\ \forall i = 1...43$. Following Morley and Piger (2012), we compute the SBIC using the specification in Davidson and MacKinnon (2004):

$$SBIC_i = -\frac{N_i}{2} log(\sum_{t=1}^{N_i} (y_t - \hat{y}_{i,t})^2) - \frac{k_i}{2} log(N_i) \tag{5}$$

where $N$ is the number of observations and $k$ is the number of parameters. For our BMA forecast, we consider AR models of orders one through twelve:

$$y_t = \alpha + \sum_{i=1}^{p} \rho_i y_{t-i} + \varepsilon_t,\ p = 1..12 \tag{6}$$

We also weight over forecasts from moving-average (MA) models of orders one through twelve:

$$y_t = \alpha + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t,\ q = 1..12 \tag{7}$$

We also weight over autoregressive moving-average (ARMA) models, again with orders one through twelve, where the number of AR components equals the number of MA components:

$$y_t = \alpha + \sum_{i=1}^{p} \rho_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t,\ p = 1..12,\ q = 1..12,\ p = q \tag{8}$$

Together, equations (6) through (8) encompass a variety of specifications of basic AR, MA, and ARMA models that might characterize a forecasted series. Beyond these three types of models, we consider two simple specifications of an unobserved components model, as described by Harvey (1989). Both specifications assume a first-order cyclical component and exclude a trend component. While the Bureau of Economic Analysis (BEA) seasonally adjusts quarterly NIPA series, some residual seasonal variation may remain, so we try versions with and without a quarterly seasonal component. The specification with a seasonal component is:

$$y_t = \sum_{j=1}^{s-1} \gamma_{t-j} + \psi_t + \varepsilon_t,\ s = 4 \tag{9}$$

We consider several specifications of autoregressive conditional heteroskedasticity (ARCH) and generalized autoregressive conditional heteroskedasticity (GARCH) models, with different lags, autoregressive terms, and in-mean ARCH terms:

$$y_t = \alpha + \sum_{i=1}^{p} \rho_i y_{t-i} + \sum_{j=0}^{q} \phi_j \sigma_{t-j}^2 + \varepsilon_t \tag{10}$$

where the error term of equation (10) contains ARCH and GARCH terms:

$$Var(\varepsilon_t) = \gamma_0 + \sum_{k=1}^{r} \alpha_{1,k} \varepsilon_{t-k}^2 + \sum_{m=1}^{s} \alpha_{2,m} \sigma_{t-m}^2 \tag{11}$$

We specify three variants of equations (10) and (11): $p = 4, q = 2, r = 2, s = 2$; $p = 0, q = 2, r = 2, s = 1$; and $p = 1, q = 1, r = 1, s = 1$.

The final set of models contains single- and double-exponential smoothed forecasts (Chatfield, 2001). The single-exponential forecast estimates a model with a single smoothing parameter:

$$\hat{y}_t = \alpha y_{t-1} + (1-\alpha)\hat{y}_{t-1} \tag{12}$$

The double-exponential version applies the smoothing process of equation (12)'s once-smoothed series.

We estimate all models with maximum likelihood. We exclude model forecasts from the BMA forecast that give an implied annual growth rate outside of [-400%, 400%] to avoid skewing the BMA forecast towards any one specific model.

# 4 Results

## 4.1 RMSEs and MAEs Using End-of-Sample Vintage Data, Fixed Sample

Figures 1 and 2 plot RMSEs of Greenbook forecasts and our forecasting benchmarks estimated on EOS vintage data, with the sample starting in 1986 Q1. We normalize the forecasting RMSEs relative to the random-walk RMSEs following Hyndman and Koehler (2006), so a number greater than one indicates a RMSE worse than the random walk. The horizontal axis denotes the forecast horizon relative to the quarter the forecast is made, so $t = 0$ indicates a forecast of the current quarter.

The Greenbook forecasts tend to significantly outperform all four of our benchmark forecasts for $t = -1$ and $t = 0$. This result is consistent with earlier evidence that Federal Reserve Board forecasters take considerable lengths to replicate the procedures of national statistical agencies' upcoming data releases (Faust and Wright, 2009; Baghestani, 2011).

However, even by the one-quarter ahead forecast horizon ($t = 1$), the relative forecast performance of the Greenbook decreases substantially. The Greenbook forecasts of government spending are comparable with the AR(1), historical mean, and BMA forecasting benchmarks at $t = 1$ while Greenbook forecasts of PCE perform comparably with the BMA forecasting benchmark at $t = 1$ but outperform the AR(1), random walk, and historical mean. The Greenbook forecast accuracy of GDP is comparable to our benchmarks by $t = 2$.

By the one-year forecast horizon ($t = 4$), the RMSEs of the Greenbook forecasts are comparable with at least one of our benchmarks for all sectors except the unemployment rate, where the Greenbook forecasts tend to do quite well against our benchmarks for the entire eight quarter forecast horizon. Greenbook forecasts of core PCE inflation are outperformed by the AR(1) and BMA for forecast horizons greater than one year.

Figures 3 and 4 display MAEs of Greenbook forecasts and our forecasting benchmarks, again estimated with EOS data with the sample starting in 1986 Q1. The results are largely similar to the RMSE results in Figures 1 and 2. The Greenbook outperforms the benchmarks in the short term across all sectors, but by the one-year forecast horizon the performance is comparable between the Greenbook and our benchmarks. For horizons longer than one year the Greenbook core PCE inflation forecasts are outperformed by our benchmarks.[7]

## 4.2 RMSEs Using Real-Time Vintage Data, Fixed Sample

Figures 5 and 6 show the results with our forecasting benchmarks estimated with RTV data, as opposed to EOS data. For GDP and its components, we use a time-series of BEA third-release (twice-revised) estimates as our RTV data whenever possible.[8] The RTV results are similar when we use a time-series of BEA first-release (never revised) or second-release (once-revised) data. For core PCE inflation and the unemployment rate, our RTV series is quarterly averages computed when the last month of data for the quarter is first available.

The results from Figures 5 and 6 are largely similar to the results estimated on EOS data. For GDP and its components, the use of RTV data tends to worsen the RMSE of the BMA forecast relative to the random walk. However, for core PCE inflation RTV data improves the BMA forecast relative to the random walk. The RMSEs of Greenbook forecasts are still comparable with at least one of our benchmarks by the one-year horizon except the unemployment rate, which still outperforms our benchmarks for the entire eight quarter forecast horizon.

[7] MAE results are always similar to the RMSE results, so we omit MAE results for the remainder of this paper.

[8] For these series, when we forecast a horizon that uses NIPA data from $t = -1$ when a third release estimate for $t = -1$ is unavailable, we use the latest available NIPA release for $t = -1$.

## 4.3 RMSEs Using End-of-Sample Vintage Data, Rolling Sample

The results so far all use a fixed sample start date of the first quarter of 1986. Figures 7 and 8 reestimate our benchmarks using a 40-quarter rolling sample, as opposed to the fixed sample start date. For example, the forecast for 1997 Q1 uses models estimated on data from 1986 Q1 to 1996 Q4. The forecast for 1997 Q2 uses models estimated on data from 1986 Q2 to 1997 Q1, and so on. The results in Figures 7 and 8 are largely similar to Figures 1 and 2, although the BMA forecasts tends to perform more poorly at longer horizons using the 40-quarter rolling sample.

## 4.4 MDE Using End-of-Sample Vintage Data, Fixed Sample

As a final method of evaluation, we check to see whether Greenbooks accurately forecast differences in growth rates, following Baghestani (2011). For each forecast horizon, we compute the difference in growth rates relative to $t = -2$, which is both the first quarter where we have a third-release NIPA series in every Greenbook and the most recent quarter for which the Federal Reserve staff does not prepare a Greenbook forecast. We define mean directional error (MDE) as the proportion of forecasts where the sign of the predicted difference in growth rates was incorrect, so higher values of MDE indicate worse forecast performance.[9]

Figures 9 and 10 show our results from the MDE measure of accuracy. We normalize the MDE relative to the MDE for the historical mean and omit the random walk, because the random walk forecast always implies no acceleration or deceleration.[10]

Figures 9 and 10 confirm that the Greenbook MDE outperforms our forecasting benchmarks for very short horizons. Much like the results for the RMSEs and MAEs, by approximately the one-year horizon, typically Greenbook MDEs are comparable with at least one of our forecasting benchmarks. The Greenbook MDEs for core PCE inflation, exports, change in business inventories, government spending, and residential investment are worse than at least one benchmark by

[9] The unemployment rate is still in levels and change in business inventories is still in first-differences, so we compute first-differences and second-differences, respectively, for these two categories.

[10] For forecasts in growth rates, the random walk predicts constant growth (no acceleration or deceleration of the growth rate), so the implied effect on the level is an acceleration.

the two-quarter forecast horizon. Notably, the Greenbook MDEs for nonresidential investment, PCE, and imports compare quite favorably to our benchmarks.

# 5 Conclusion

This paper compares Greenbook forecasts of the unemployment rate, core PCE inflation, GDP, and the major components of GDP to forecasts from several simple univariate benchmarks. The primary contribution of this paper is to analyze a wider range of Greenbook projections in a unified framework than previous studies.

We find that Greenbook forecasts generally outperform our simple benchmarks in the very short forecasting horizon. However, typically by the one-year forecast horizon, the accuracy of Greenbook forecasts is comparable with or worse than at least one of our benchmarks. These results hold whether we measure forecast accuracy as RMSE, MAE, or MDE. These results are consistent with earlier evidence that Greenbook forecasts carefully attempt to replicate the data release procedures of US statistical agencies, such as taking into account leading indicators that the BEA uses in constructing GDP, which gives Greenbook forecasts high short-term accuracy (Faust and Wright, 2009).

One caveat to our analysis is that we treat Greenbook forecasts as unconditional forecasts, comparing the forecasts to the unconditional forecasts generated from our benchmark models. However, in practice the Greenbook forecasts are conditioned on an exogenous path for policy. Explicitly taking into account the conditional nature of Greenbook forecasts may either improve or diminish their measured accuracy.

# References

Arai, Natsuki, "Using Forecast Evaluation to Improve the Accuracy of the Greenbook Forecast," International Journal of Forecasting 30:1 (2014), 12-19.

Baghestani, Hamid, "Federal Reserve and Private Forecasts of Growth in Investment," Journal of Economics and Business 63:4 (2011), 290-305.

Chatfield, Chris, Time-Series Forecasting, London: Chapman & Hall/CRC (2001).

Clements, Michael P., and Ana Beatriz Galvão, "Real-Time Forecasting of Inflation and Output Growth With Autoregressive Models in the Presence of Data Revisions," Journal of Applied Econometrics 28:3 (2013), 458-477.

Davidson, Russell, and James G. MacKinnon, Econometric Theory and Methods, New York: Oxford University Press (2004).

Faust, Jon, and Jonathan H. Wright, "Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset," Journal of Business & Economic Statistics 27:4 (2009), 468-479.

Federal Reserve Bank of St. Louis. ALFRED. alfred.stlouisfed.org. Various Access Dates.

Federal Reserve Board. Federal Open Market Committee. Transcripts and Other Historical Materials. www.federalreserve.gov/monetarypolicy/fomc_historical.htm. Various Access Dates.

Harvey, Andrew C., Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge: Cambridge University Press (1989).

Hyndman, Rob J., and Anne B. Koehler, "Another Look at Measures of Forecast Accuracy," International Journal of Forecasting 22:4 (2006), 679-688.

Koenig, Evan F., Sheila Dolmas, and Jeremy Piger, "The Use and Abuse of Real-Time Data in Economic Forecasting," Review of Economics and Statistics 85:3 (2003), 618-628.

Morley, James, and Jeremy Piger, "The Asymmetric Business Cycle," Review of Economics and Statistics 94:1 (2012), 208-221.

Raftery, Adrian E., “Bayesian Model Selection in Social Research,” Sociological Methodology 25 (1995), 111-163.

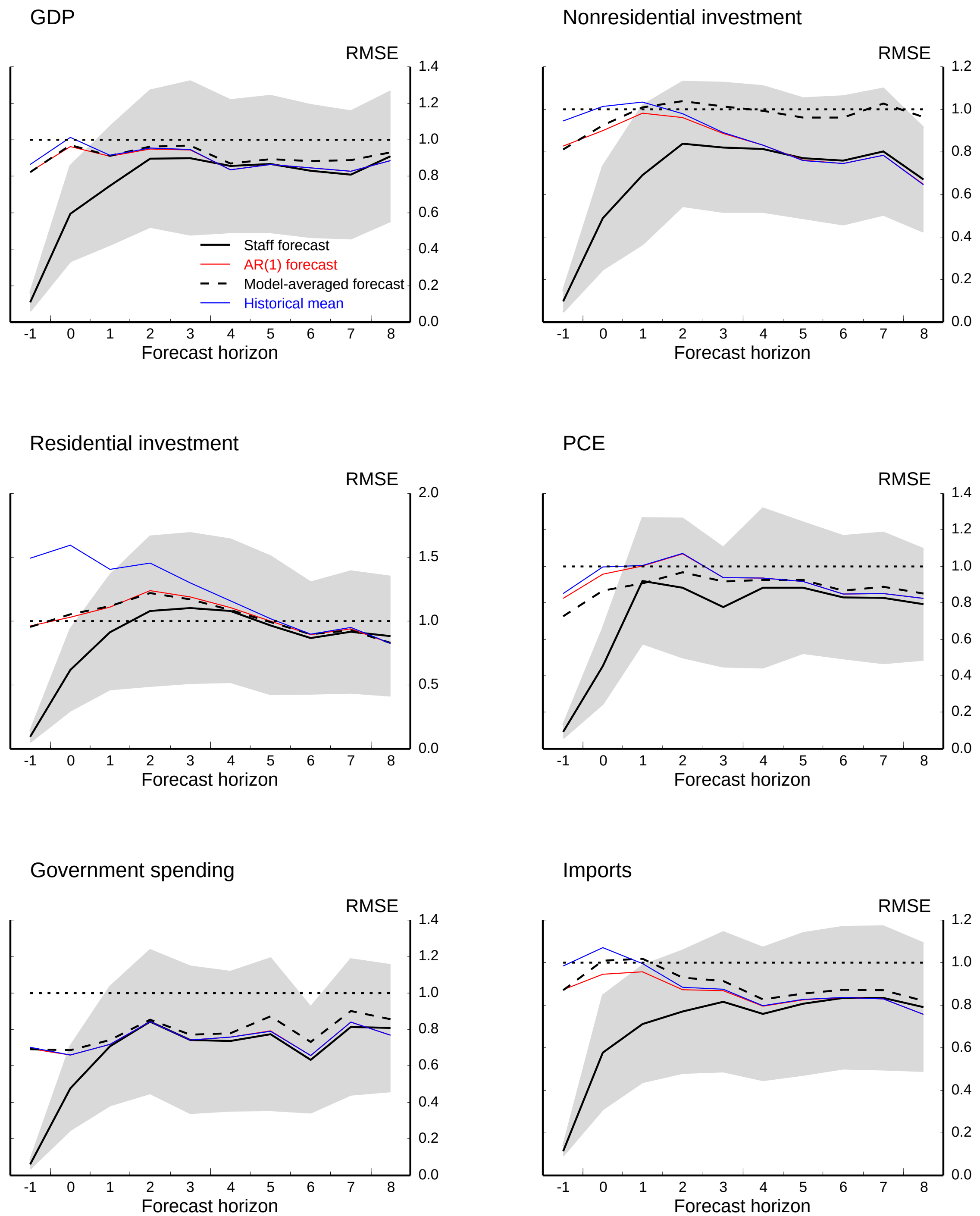
Reifschneider, David, and Peter Tulip, “Gauging the Uncertainty of the Economic Outlook From Historical Forecasting Errors,” FEDS Working Paper 2007-60 (2007).

Romer, Christina D., and David H. Romer, “Federal Reserve Information and the Behavior of Interest Rates,” American Economic Review 90:3 (2000), 429-457.

Tulip, Peter, “Has the Economy Become More Predictable? Changes in Greenbook Forecast Accuracy”, Journal of Money, Credit and Banking 41:6 (2009), 1217-1231.

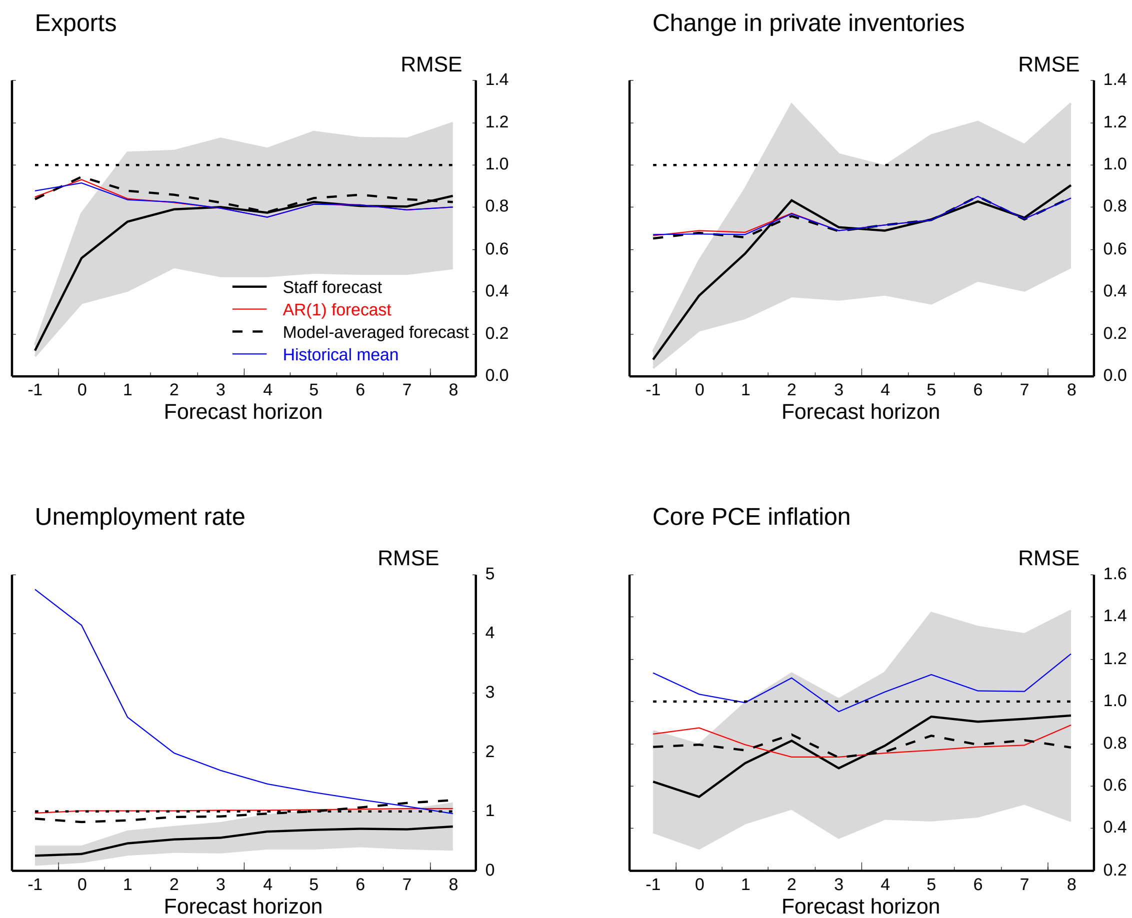
Wright, Jonathan H., “Forecasting US Inflation by Bayesian Model Averaging,” Journal of Forecasting 28:2 (2009), 131-144.

Figure 1: RMSEs Using End-of-Sample Vintages, Fixed Sample



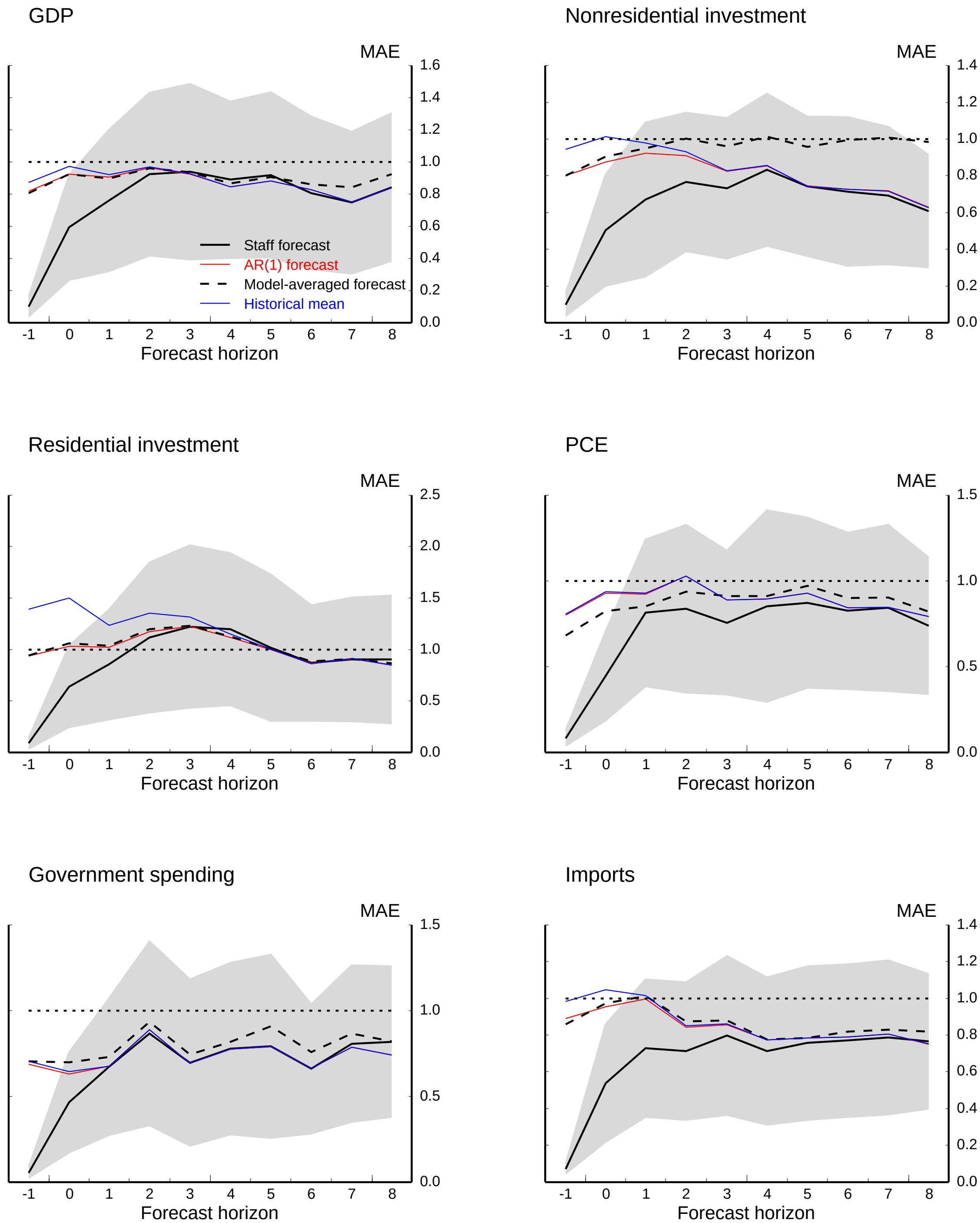

Root mean squared errors (RMSEs) normalized relative to the RMSE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the RMSEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 2: RMSEs Using End-of-Sample Vintages, Fixed Sample (Continued)



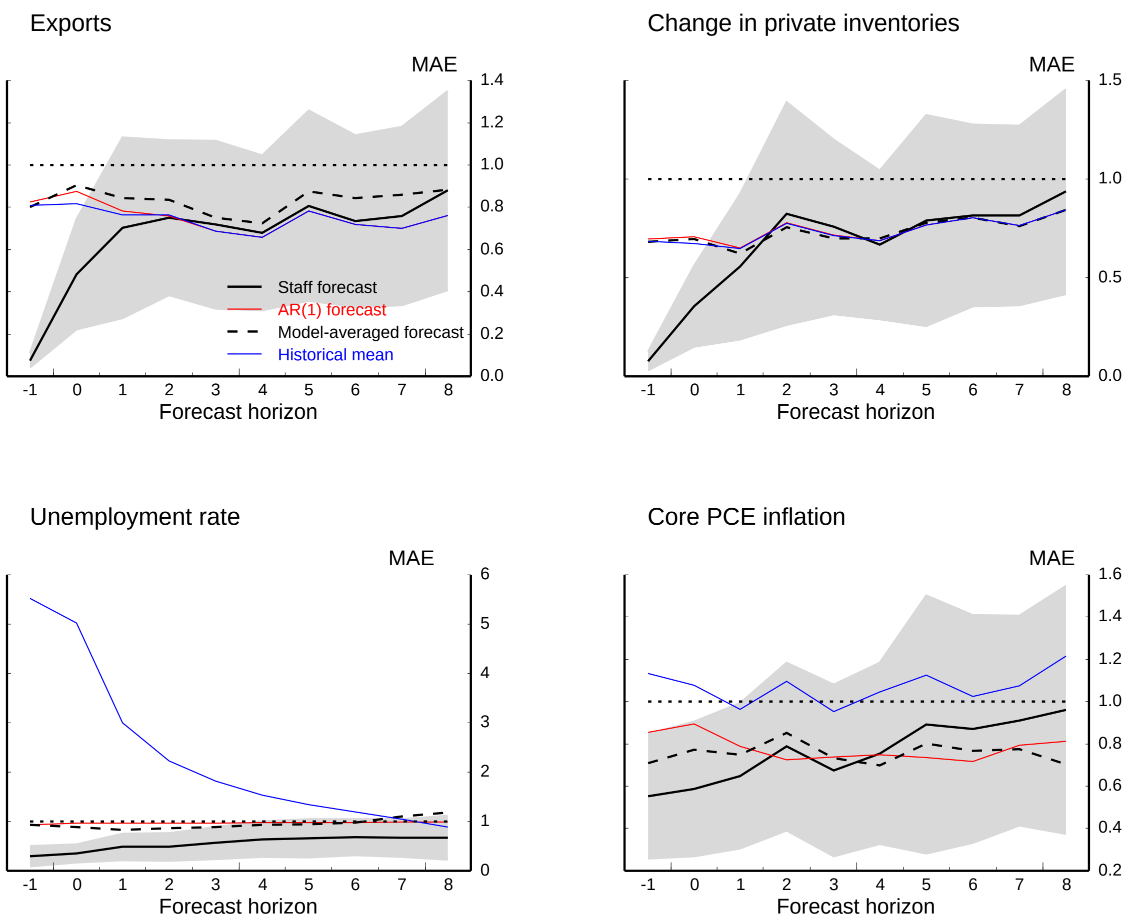

Root mean squared errors (RMSEs) normalized relative to the RMSE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the RMSEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 3: MAEs Using End-of-Sample Vintages, Fixed Sample



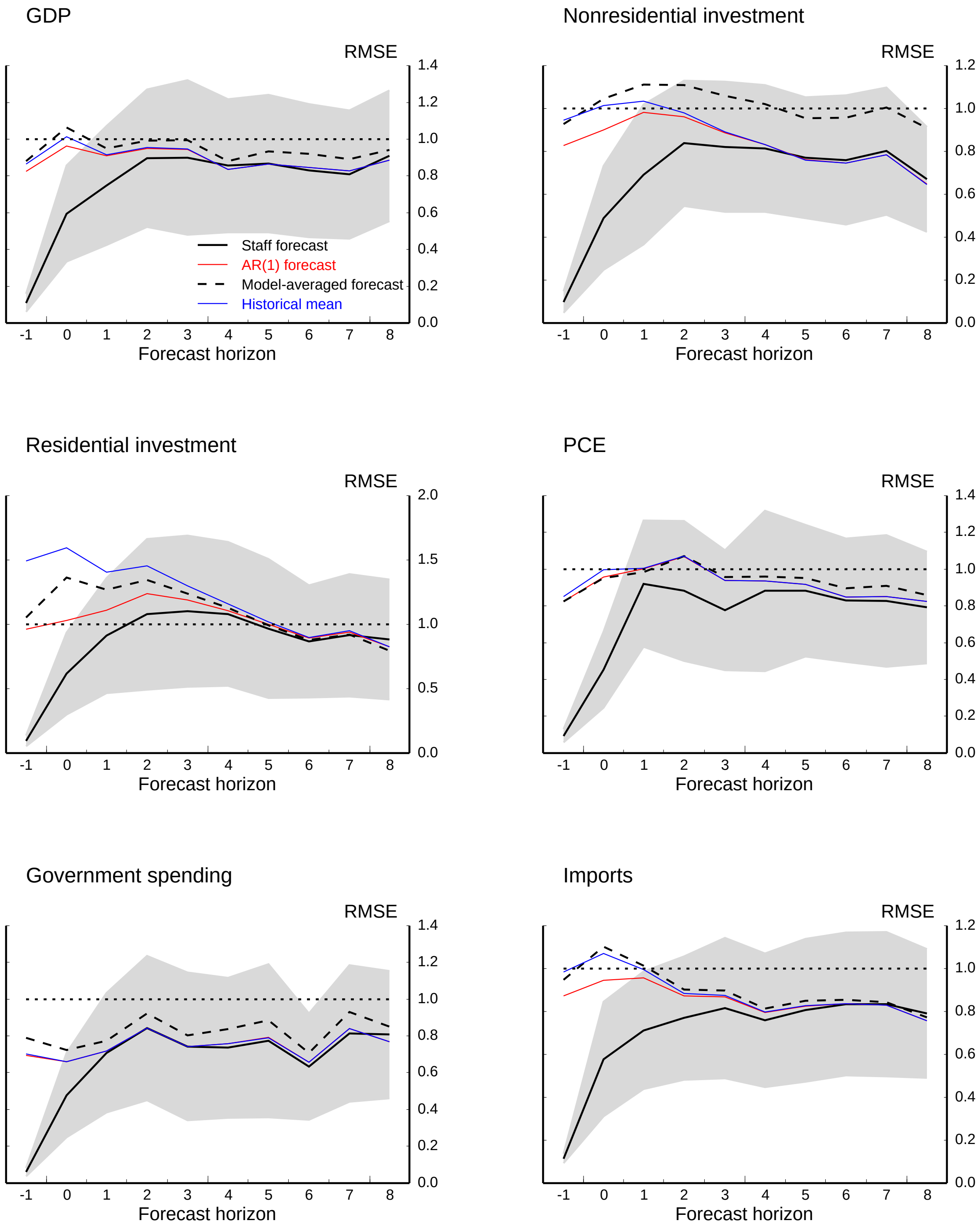

Mean absolute errors (MAEs) normalized relative to the MAE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the MAEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 4: MAEs Using End-of-Sample Vintages, Fixed Sample (Continued)



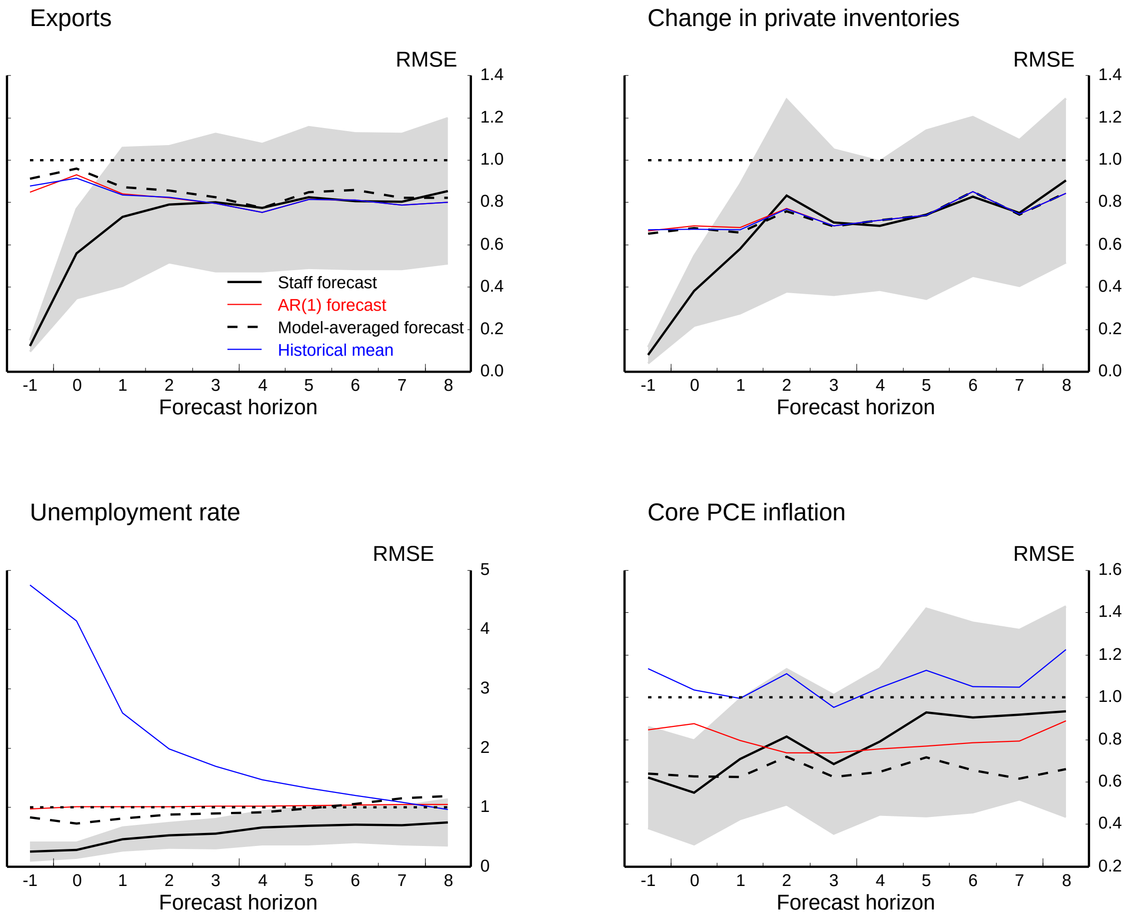

Mean absolute errors (MAEs) normalized relative to the MAE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the MAEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 5: RMSEs Using Real-Time Vintages, Fixed Sample



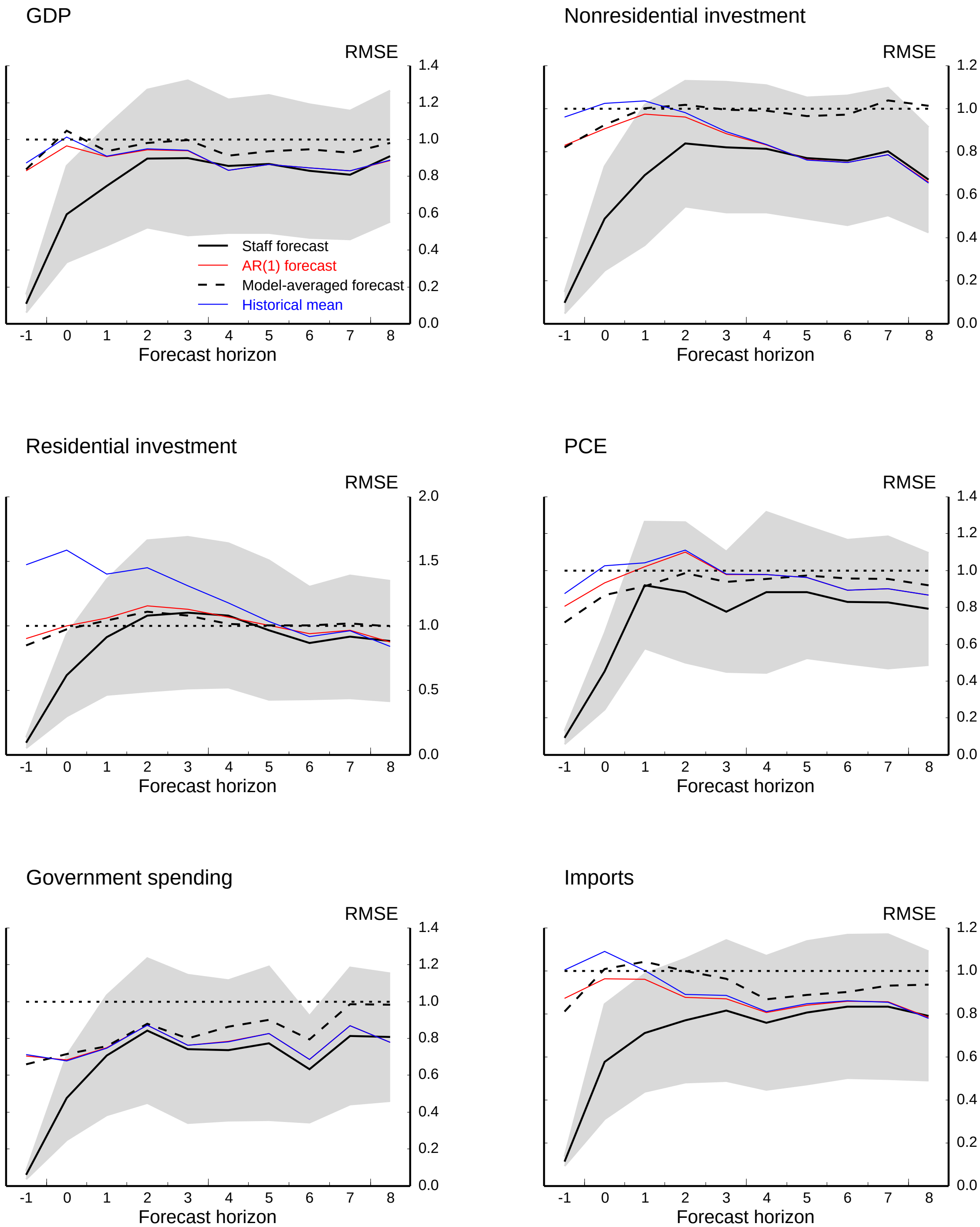

Root mean squared errors (RMSEs) normalized relative to the RMSE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the RMSEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 6: RMSEs Using Real-Time Vintages, Fixed Sample (Continued)



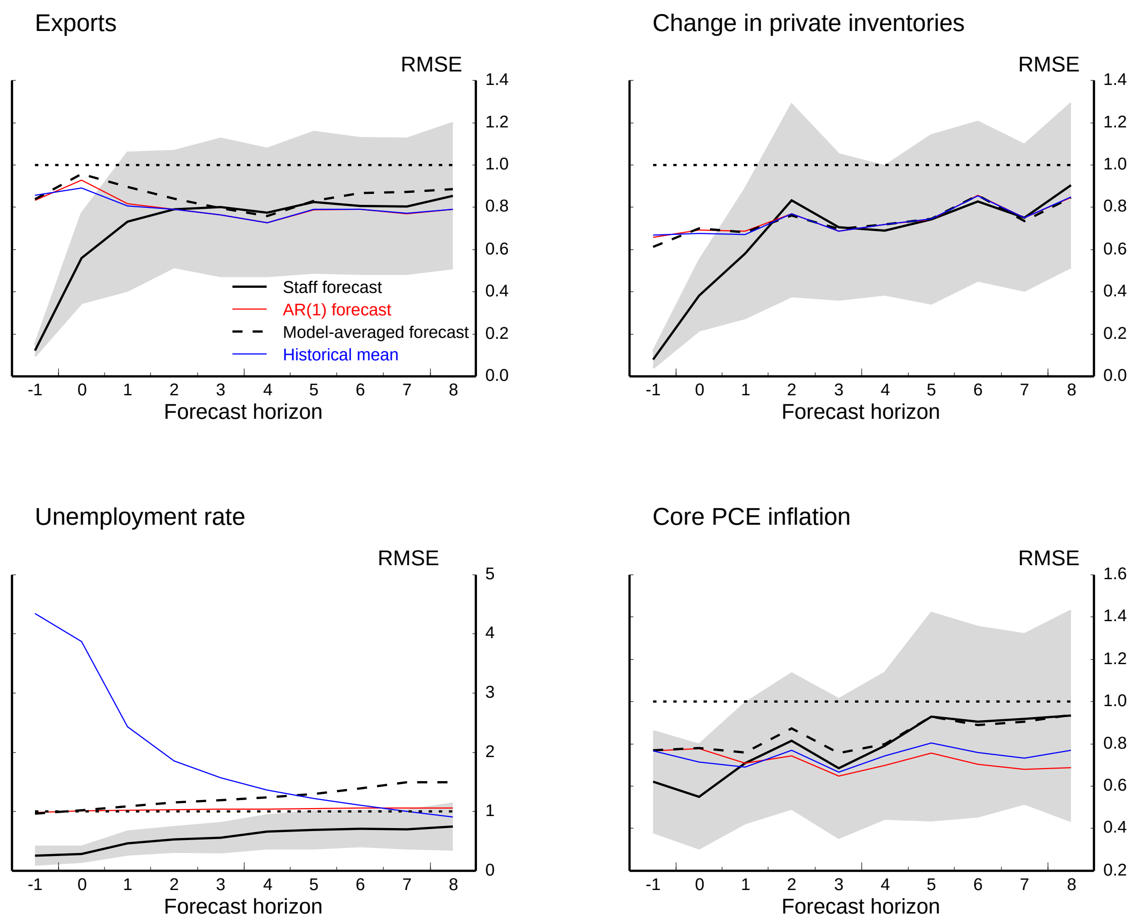

Root mean squared errors (RMSEs) normalized relative to the RMSE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the RMSEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 7: RMSEs Using End-of-Sample Vintages, 40-Quarter Rolling Sample



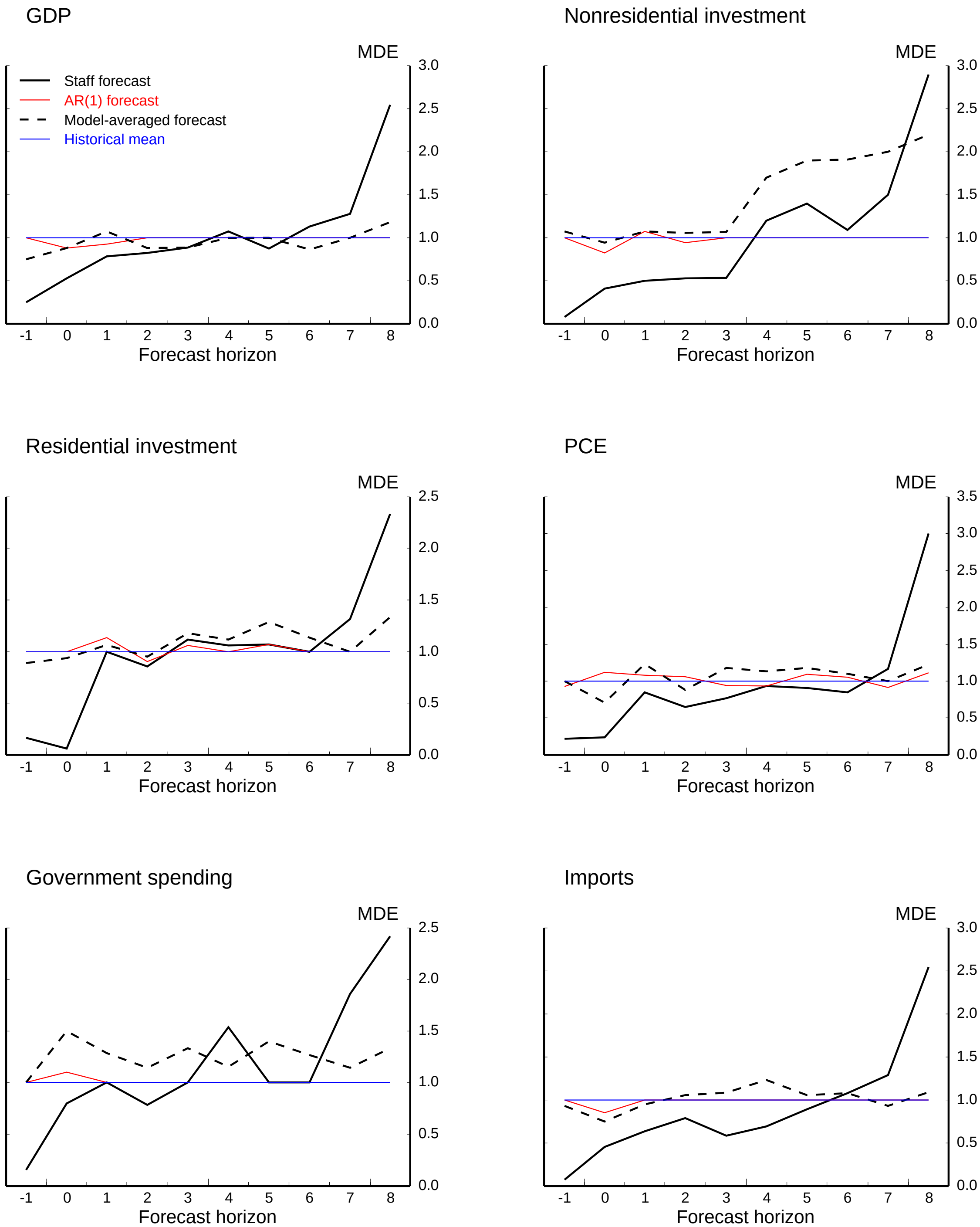

Root mean squared errors (RMSEs) normalized relative to the RMSE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the RMSEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 8: RMSEs Using End-of-Sample Vintages, 40-Quarter Rolling Sample (Continued)



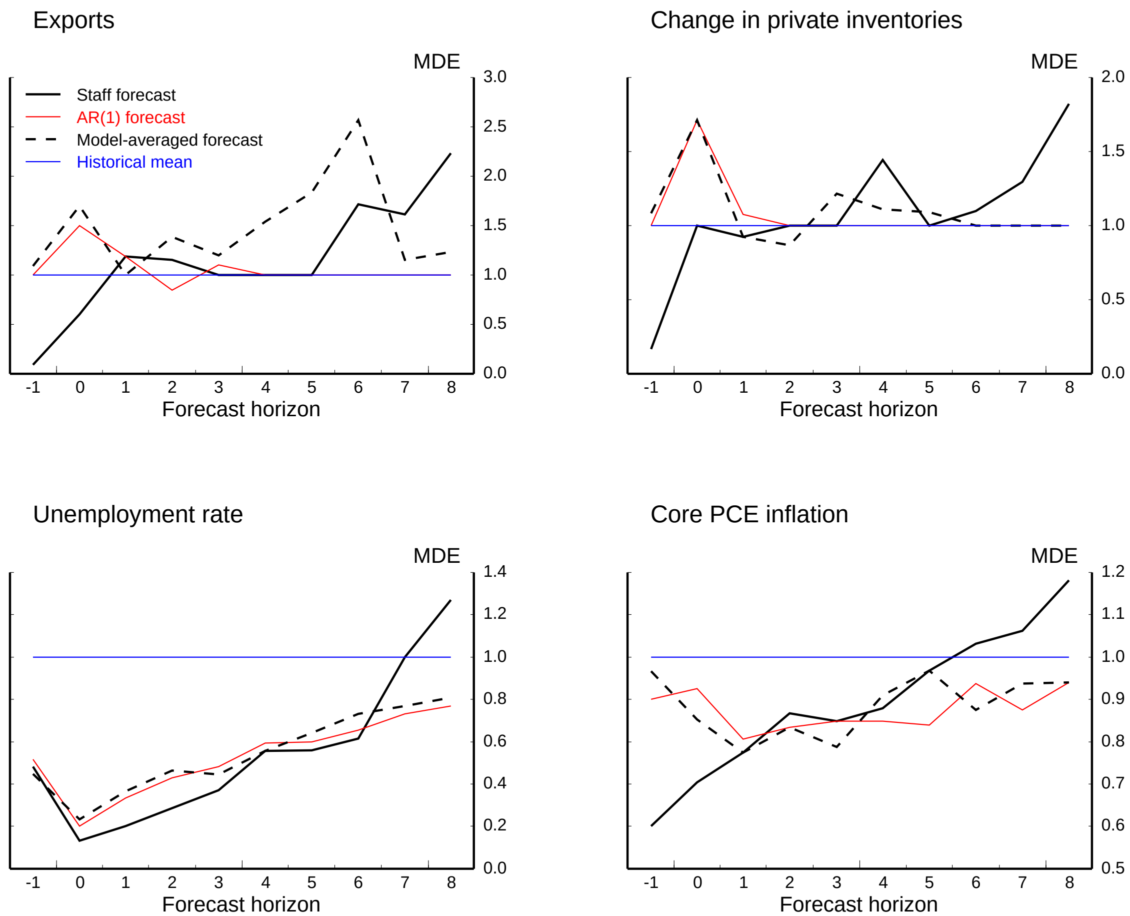

Root mean squared errors (RMSEs) normalized relative to the RMSE for the random walk for each forecast horizon-sector (Hyndman and Koehler, 2006). The horizontal axis denotes the horizon relative to when the forecast is made. Shaded area indicates the RMSEs of the Greenbook forecast calculated with the 68th percentile of Greenbook forecast errors, centered on the Greenbook forecast.

Figure 9: MDEs Using End-of-Sample Vintages, Fixed Sample




Mean directional errors (MDEs) normalized relative to the MDE for the historical mean for each forecast horizon-sector. The horizontal axis denotes the horizon relative to when the forecast is made.

Figure 10: MDEs Using End-of-Sample Vintages, Fixed Sample (Continued)




Mean directional errors (MDEs) normalized relative to the MDE for the historical mean for each forecast horizon-sector. The horizontal axis denotes the horizon relative to when the forecast is made.

Table 1: Summary Statistics

| Vintage of Data | Variable | Mean | Standard Deviation | Minimum | Maximum |
|---|---|---|---|---|---|
| End-of-sample, December 2008 Greenbook | GDP | 2.89 | 2.06 | -2.98 | 7.48 |
| | Consumption | 3.12 | 1.99 | -3.74 | 7.12 |
| | Nonresidential Investment | 4.55 | 7.95 | -13.57 | 22.10 |
| | Residential Investment | 1.31 | 11.55 | -27.01 | 24.15 |
| | Government Spending | 2.01 | 3.20 | -4.48 | 9.43 |
| | Exports | 7.31 | 7.91 | -18.18 | 27.36 |
| | Imports | 6.45 | 7.37 | -12.62 | 18.04 |
| | Business Inventories | -0.71 | 29.74 | -76.60 | 76.50 |
| | Unemployment Rate | 5.53 | 0.94 | 3.90 | 7.63 |
| | Core PCE Inflation | 2.53 | 1.08 | 0.84 | 5.37 |
| Real-time | GDP | 2.99 | 2.09 | -2.79 | 8.20 |
| | Consumption | 3.17 | 1.99 | -3.39 | 7.74 |
| | Nonresidential Investment | 5.99 | 8.74 | -16.30 | 26.74 |
| | Residential Investment | 1.77 | 11.77 | -25.33 | 31.66 |
| | Government Spending | 1.74 | 3.91 | -8.31 | 16.68 |
| | Exports | 7.33 | 8.94 | -18.76 | 30.79 |
| | Imports | 7.21 | 8.10 | -12.99 | 22.31 |
| | Business Inventories | -1.33 | 27.02 | -74.90 | 91.60 |
| | Unemployment Rate | 5.53 | 0.94 | 3.90 | 7.63 |
| | Core PCE Inflation | 2.59 | 1.10 | 1.05 | 5.36 |

Unemployment rate is in levels. Inventories is in real first-differenced billions of dollars seasonally adjusted at an annual rate. All other variables are in real seasonally adjusted annual percent changes. Real-time data for gross domestic product (GDP) and its components are the BEA's third-release twice-revised ("final") release as of the December 2008 Greenbook. Real-time data for the unemployment rate and core personal consumption expenditures (PCE) inflation are quarterly averages computed when the last month of data for the quarter is first available.