# Supplemental Appendix to "Modeling Time-Varying Uncertainty of Multiple-Horizon Forecast Errors"*

Todd E. Clark
Federal Reserve Bank of Cleveland
todd.clark@clev.frb.org

Michael W. McCracken
Federal Reserve Bank of St. Louis
michael.w.mccracken@stls.frb.org

Elmar Mertens
Bank for International Settlements
elmar.mertens@bis.org

May 2018

## Contents



*The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, Federal Reserve Bank of St. Louis, Federal Reserve System, or the Bank for International Settlements.

# List of Figures

# List of Tables

# I References to Examples of Central Bank Fan Charts

Table A.1 provides links to fan chart documents for the countries we mention: the Reserve Bank of Australia (RBA), the European Central Bank (ECB), the Federal Reserve (Fed), and the Bank of England (BoE).

Table A.1: References: Central bank fan charts

- RBA fan charts, February 2018: https://www.rba.gov.au/publications/smp/2018/feb/pdf/06-economic-outlook.pdf
- RBA documents explaining the construction of the charts: https://www.rba.gov.au/speeches/2016/sp-ag-2016-04-06.html#fn9 and https://www.rba.gov.au/publications/rdp/2012/pdf/rdp2012-07.pdf
- ECB fan charts, March 2018: https://www.ecb.europa.eu/pub/pdf/other/ecb.ecbstaffprojections201803.en.pdf?ed72891ff06f1793746fd5704d13595c
- ECB document explaining the construction of the charts: https://www.ecb.europa.eu/pub/pdf/other/newprocedureforprojections200912en.pdf?917c588c190f66c694c463d2cee934d4
- Fed fan charts, December 2017: https://www.federalreserve.gov/monetarypolicy/fomcminutes20170315ep.htm
- Fed documentation explaining the construction of the charts: see the notes to Figures 4.A, 4.B, 4.C, and 5, and https://www.federalreserve.gov/econresdata/feds/2017/files/2017020pap.pdf
- BoE fan charts, February 2018: https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2018/february/inflation-report-february-2018.pdf?la=en&hash=555ED88EF574D368B81BF703480C1987EEBBA883
- BoE documents explaining the construction of the charts: https://www.bankofengland.co.uk/-/media/boe/files/inflation-report/2002/may-2002.pdf?la=en&hash=9E8BD9AA0E9D95CA8D95F66569A61B318D189DA3

# II Expectational Updates in General Forecasting Framework

This section shows that our proposed multi-horizon stochastic volatility (SV) specification for historical forecast errors from a source such as the SPF can be seen as having a basis in a general time series forecasting model. We stress that we don't take the SPF forecasts as literally or directly coming from such a model, and, as we note in the paper, we take the forecasts and errors as given rather than try to improve them. In summarizing a 2009 survey of the SPF participants, Stark (2013) indicates panelists reported using models to forecast and applying subjective judgment to adjust the model-based forecasts. They also indicated that the role of models changes with the forecast horizon. Accordingly, our intention is to establish a broad basis for our proposed multi-horizon SV in a general time series model without taking the model as a "true" specification underlying all of the SPF respondents' forecasts, such that the general model were used to establish restrictions on a multi-horizon SV model applied to the SPF data on expectational updates.

In this treatment, for simplicity, we abstract from nowcasts, which arise in forecasting with real-time data but not simple time series models; our treatment here is easily extended to include nowcasts. As detailed in the paper, our basic model in expectational updates takes the form:

$$\boldsymbol{\eta}_t = (E_t - E_{t-1}) \begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+H-1} \end{bmatrix} = \boldsymbol{A}\boldsymbol{\Lambda}_t^{0.5}\boldsymbol{\varepsilon}_t \qquad \varepsilon_t \sim N(\mathbf{0}, \boldsymbol{I}) \tag{A.1}$$

where $\boldsymbol{A}$ is unit-lower-triangular, and $\boldsymbol{\eta}_t$ and $\boldsymbol{\varepsilon}_t$ have the same length.

## II.1 General time series forecasting model

To consider a more general basis for our model in expectational updates, suppose that the dynamics for the scalar variable $y_t$ of interest can be characterized by a linear state-space model with state vector $\boldsymbol{X}_t$ that includes $y_t$ as one of its variables:

$$\boldsymbol{X}_t = \boldsymbol{F}\boldsymbol{X}_{t-1} + \boldsymbol{w}_t \qquad E_{t-1}\boldsymbol{w}_t = 0 \tag{A.2}$$

$$y_t = \boldsymbol{g}\,\boldsymbol{X}_t, \tag{A.3}$$

where $\boldsymbol{g}$ is a row vector that selects $y_t$. For the moment, we treat the shocks $\boldsymbol{w}_t$ to the state-vector merely as martingale differences; below we will add assumptions on their second moments.

In this formulation, the transition equation in the states $\boldsymbol{X}_t$ captures an array of time series specifications, including VARs with lag order greater than one (and written in equation (A.2) in companion form). We think of $\boldsymbol{X}_t$ as containing the observed variables of a potentially large VAR capturing macroeconomic dynamics; although we treat everything as observed, unobserved components could be included in $\boldsymbol{X}_t$ without change to our main points.

This dynamic specification implies that forecasts and forecast updates are given by:

$$E_t\boldsymbol{X}_{t+h} = \boldsymbol{F}^h\boldsymbol{X}_t \tag{A.4}$$

$$(E_t - E_{t-1})\boldsymbol{X}_{t+h} = \boldsymbol{F}^h\boldsymbol{w}_t \tag{A.5}$$

$$(E_t - E_{t-1})y_{t+h} = \boldsymbol{g}\,\boldsymbol{F}^h\boldsymbol{w}_t, \tag{A.6}$$

which yields a vector of forecast updates that corresponds to a linear combination of the state-vector innovations $\boldsymbol{w}_t$:

$$\boldsymbol{\eta}_t = \begin{bmatrix} \boldsymbol{g} \\ \boldsymbol{g}\,\boldsymbol{F} \\ \vdots \\ \boldsymbol{g}\,\boldsymbol{F}^{H-1} \end{bmatrix} \boldsymbol{w}_t = \boldsymbol{D}\boldsymbol{w}_t. \tag{A.7}$$

The coefficients of the linear combinations are functions of the coefficients of the transition equation's coefficient matrix $\boldsymbol{F}$. With $\boldsymbol{g}$ selecting variable $j$ of $\boldsymbol{X}_t$, rows 2 through $H-1$ of the matrix $\boldsymbol{D}$ contain, respectively, the $j$'th row of $\boldsymbol{F}$, $\boldsymbol{F}^2$, ..., $\boldsymbol{F}^{H-1}$.

The specific implications of the formulation (A.7) depend in part on the dimensions of the state-vector innovations $\boldsymbol{w}_t$ (denoted $N_w$) and the selected vector of expectational updates (denoted $N_\eta$). There is no reason that these dimensions should be the same. The former can be thought of as mirroring the number of (stochastic) signals observed by the forecaster, whereas the latter merely reflects the number of forecast horizons considered. In the SPF data, although we observe commonality in the observed expectational updates, we do not observe a singular variance-covariance matrix for $\boldsymbol{\eta}_t$. Accordingly, we can assume that $N_w \geq N_\eta$.

But as we noted in an earlier version of this paper,[1] if — counterfactually, as regards the rank of the sample estimate of $\text{Var}(\boldsymbol{\eta}_t)$ — the data are assumed to be governed by a univariate AR process, the simple time series specification implies restrictions on the model of expectational updates. In particular, the vector of expectational updates would be a function of a single scalar shock $w_t$. In our earlier draft, in discussing the implications of an AR-SV model for $y_t$, we overstated the extent to which a simple time series model implies common factor restrictions for the expectational updates. Rather, as this more general time series formulation shows, with more than one variable, the single-factor restriction we previously discussed no longer applies.

We now turn to the implications of the general time series model, first in the case of constant innovation variances and then in the case of stochastic volatility.

[1]See `https://ideas.repec.org/p/fip/fedcwp/1715.html`.

## II.2 Constant variance case

Consider first our multiple-horizon model of expectational updates modified to treat variances as constant over time:

$$\boldsymbol{\eta}_t = \boldsymbol{A}\boldsymbol{\Lambda}^{0.5}\boldsymbol{\varepsilon}_t, \tag{A.8}$$

where $\boldsymbol{\varepsilon}_t$ is a standard-normal vector of same length as $\boldsymbol{\eta}_t$, $\boldsymbol{A}$ is a unit-lower-triangular matrix, and $\boldsymbol{\Lambda}$ is diagonal. This formulation implies:

$$\boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{A}\boldsymbol{\Lambda}\boldsymbol{A}'). \tag{A.9}$$

Similarly, for the general time series model, suppose that the variance of the innovation vector $\boldsymbol{w}_t$ is constant over time: specifically, let $\boldsymbol{w}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. It then follows from (A.7) that

$$\boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}'). \tag{A.10}$$

The time series model-implied representation (A.10) of course resembles our specification (A.9). In fact, for our purposes they are observationally equivalent, so that if innovation variances are constant over time, our model restricted to treat $\boldsymbol{\Lambda}$ as constant is entirely consistent with a general linear time series model.

Although our proposed specification is consistent with a general time series model, we don't mean to suggest the specifications are the same. Typically, the coefficient matrix $\boldsymbol{D}$ would *not* be unit-lower triangular, and the variance-covariance matrix of $\boldsymbol{w}_t$ would be dense (rather than diagonal). So it is not the case that $\boldsymbol{A}$ would map to $\boldsymbol{D}$ or $\boldsymbol{\Lambda}$ would map to $\boldsymbol{\Sigma}$. In general, as long as $N_w > N_\eta$, we cannot recover $\boldsymbol{w}_t$ from observing $\boldsymbol{\eta}_t$, but such non-invertibility does not matter for our purposes.

## II.3 Stochastic volatility case

Now consider the case in which the general time series model features stochastic volatility. In particular, reflecting the formulation commonly used with VARs, let $\boldsymbol{w}_t = \boldsymbol{B}\boldsymbol{\Gamma}_t^{0.5}\boldsymbol{e}_t$, where $\boldsymbol{e}_t \sim N(\mathbf{0}, \boldsymbol{I}_{N_w})$, $\boldsymbol{B}$ is unit-lower triangular, and $\boldsymbol{\Gamma}_t$ is a diagonal matrix containing $N_w$ volatilities that, in logs, follow random walk processes. Accordingly, $\boldsymbol{w}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_t)$, with $\boldsymbol{\Sigma}_t = \boldsymbol{B}\boldsymbol{\Gamma}_t\boldsymbol{B}'$. It then follows from (A.7) that

$$\boldsymbol{\eta}_t = \boldsymbol{D}\boldsymbol{B}\boldsymbol{\Gamma}_t^{0.5}\boldsymbol{e}_t \sim N(\mathbf{0}, \boldsymbol{D}\boldsymbol{B}\boldsymbol{\Gamma}_t\boldsymbol{B}'\boldsymbol{D}'). \tag{A.11}$$

In this representation based on the general time series model, the component $\boldsymbol{\Gamma}_t^{0.5}\boldsymbol{e}_t$ is a multivariate stochastic volatility process. The vector of expectational updates $\boldsymbol{\eta}_t$ is then a linear combination of the underlying stochastic volatility processes.

As a result, our multiple-horizon SV formulation in (A.1) is consistent with the general time series model. Our model, as does the general time series specification, implies the vector of expectational updates to be a linear combination of underlying stochastic volatility processes. In our model, this combination is given by $\boldsymbol{\Lambda}_t^{0.5}\boldsymbol{\varepsilon}_t$.

While our proposed model is consistent with the general specification, we don't mean to treat it as the same. In general, the formulation we apply in (A.1) will not directly recover underlying processes and coefficients of the linear model corresponding to (A.11). Our model multiplies stochastic volatility processes by a unit-lower triangular matrix $\boldsymbol{A}$. The general time series model multiplies stochastic volatility processes by the matrix product $\boldsymbol{D}\boldsymbol{B}$. Although $\boldsymbol{B}$ is unit-lower triangular, the product $\boldsymbol{D}\boldsymbol{B}$ will not be unit-lower triangular, except in the special case that the coefficient matrix $\boldsymbol{F}$ is also unit-lower triangular. Again, though, given that the SPF forecasts are not purely model-based, it is not our intention to use a time series model to improve the forecasts or impose specific restrictions on our model of expectational updates of the SPF projections. Our point here is simply that our multiple-horizon SV formulation in (A.1) is consistent with a general time series model, in that the expectational updates are linear combinations of stochastic volatility processes.

## II.4 Accommodating measurement error and data revisions

The framework described above can be extended to accommodate various specifications of measurement error or data revisions. In this section, we sketch treatments of two different types of error or revisions.

First, suppose that the survey-based SPF forecasts contain some form of measurement error or noise. In this case, for the purpose of quantifying uncertainty around the SPF forecasts, we would want our estimates of uncertainty to reflect the contributions of measurement error or noise. But the general time series model of (A.2) would be unaffected by this form of measurement error.

More specifically, let $\tilde{\boldsymbol{\eta}}_t$ denote the vector of expectational updates observed in the SPF data, and let $\boldsymbol{v}_t$ denote an $N_\eta \times 1$ vector of measurement errors distributed (independently) $N(\mathbf{0}, \boldsymbol{V})$. With the measurement error,

$$\tilde{\boldsymbol{\eta}}_t = \boldsymbol{\eta}_t + \boldsymbol{v}_t. \tag{A.12}$$

For simplicity, assume all variances to be constant over time. It follows that the general time series model implies:

$$\tilde{\boldsymbol{\eta}}_t \sim N(\mathbf{0}, \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}' + \boldsymbol{V}). \tag{A.13}$$

Our proposed model specification would take the form:

$$\tilde{\boldsymbol{\eta}}_t \sim N(\mathbf{0}, \tilde{\boldsymbol{A}}\tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{A}}'), \tag{A.14}$$

with $\tilde{\boldsymbol{A}}$ unit-lower triangular.

Accordingly, for our purposes, our model of expectational updates would be observationally equivalent to the general time series model. In this constant variance case, allowing measurement error in the SPF forecasts would not break the consistency of our model with a general time series model. Allowing stochastic volatility as in our model of expectational updates and in the innovations $\boldsymbol{w}_t$ of the general time series model would be more complicated, but again, our proposed model for the SPF expectational updates would be broadly consistent with the implications of a general time series model.

Now suppose instead that the underlying macroeconomic data of interest are measured with error and subject to revision. In this case, the general time series model can be extended to include such error in a measurement equation. In particular, in the time series model along the lines of (A.2), the forecasts could be generated by an unobserved component model with latent state vector $\boldsymbol{S}_t$ in lieu of $\boldsymbol{X}_t$ and with a measurement vector $\boldsymbol{Z}_t$. Such a model is, however, observationally equivalent to the one above without any such measurement issues, with $\boldsymbol{X}_t = E(\boldsymbol{S}_t|\boldsymbol{Z}^t)$ and $\boldsymbol{w}_t = E(\boldsymbol{S}_t|\boldsymbol{Z}^t) - E(\boldsymbol{S}_t|\boldsymbol{Z}^{t-1})$ — i.e., the innovation representation of the unobserved component model.

In this case, our basic point about the consistency of our proposed model with the implications of a general time series model still applies. More general treatments of data revisions could be accommodated by extending the time series model to include the various news and noise features developed in such work as Kishor and Koenig (2012) and Jacobs, et al. (2015). Again, though, it is our intention to point out a broad basis for our model of the SPF forecast errors in a more general time series model without going so far as to take the general model as a "true" specification underlying all of the the SPF respondents' forecasts, such that the general model were used to establish restrictions on a multi-horizon SV model applied to the SPF data on expectational updates.

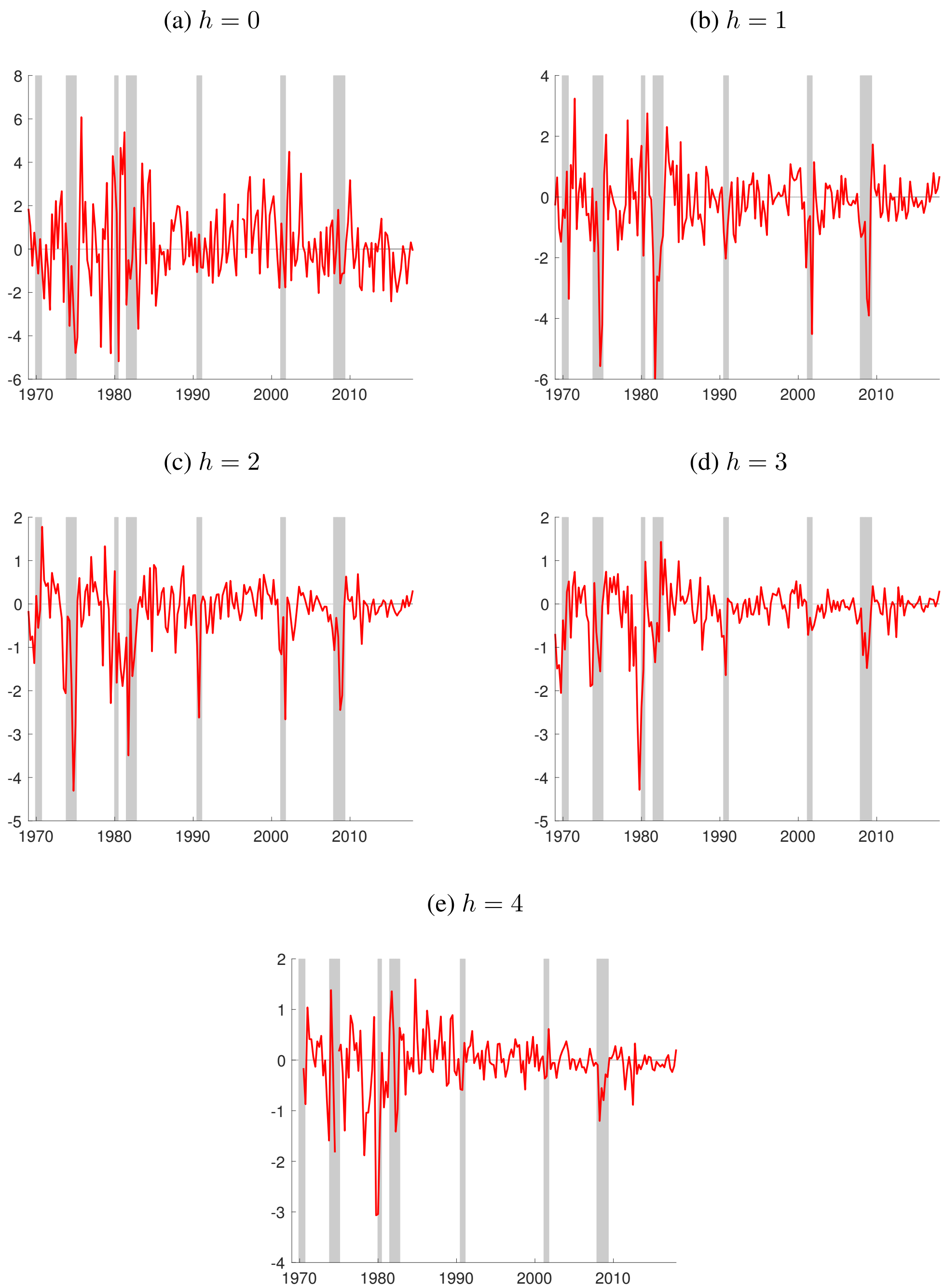
# III Additional Figures: Data and Fan Charts

Figures A.1 through A.5 display our data for expectational updates in real GDP growth, the unemployment rate, GDP price index inflation, CPI inflation and the 3-month T-bill rate (expectations based on the SPF). Figures A.6 through A.10 display the corresponding forecast error data. Section 5.1 of the paper briefly discusses some key features of the data.

Fan charts for the SPF forecasts at selected dates are illustrated in Figures A.11 through A.15. These fan charts are constructed with the same basic approach described in the paper for its charts of forecast errors and confidence bands, but centered around the forecasts rather than the forecast errors. Consistent with common central bank fan charts, these charts show the recent history of actual data, the point forecast, and confidence bands, specified as one-standard-deviation bands obtained from out-of-sample predictions generated from our baseline SV model (henceforth: "ETA-SV", red dashed lines) and the FE-CONST alternative (blue, dash-dotted lines), respectively.[2] The charts also display the outcomes realized. Across the forecast origins selected and displayed (1987:Q1, 1997:Q1, 2007:Q1, and 2017:Q1), the confidence bands in Figures A.11 through A.15 display the same patterns described in the paper's discussion of Figures 3 and 4. For example, with GDP growth, the widths of both ETA-SV and FE-CONST bands vary some across time (from one forecast origin to another), and — at the selected dates — the ETA-SV bands turn out to be generally narrower than the FE-CONST bands.

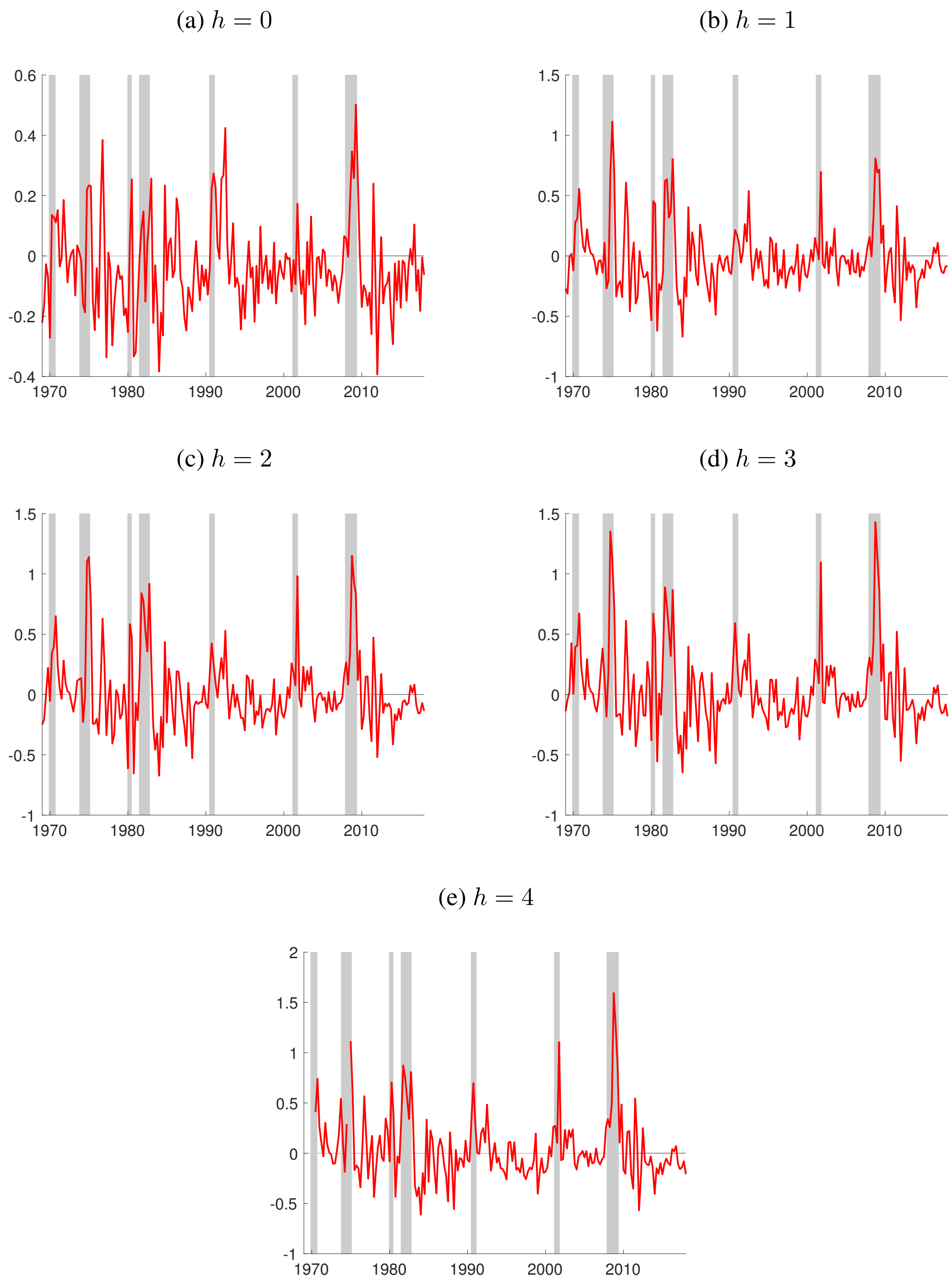
[2]Given the number of alternative models presented in this supplemental appendix, we will reference the baseline SV model as as ETA-SV rather than just SV — as we did in the main text — to emphasize that our baseline model uses $\boldsymbol{\eta}_t$ as a primitive.

Figure A.1: Expectational Updates for Real GDP Growth



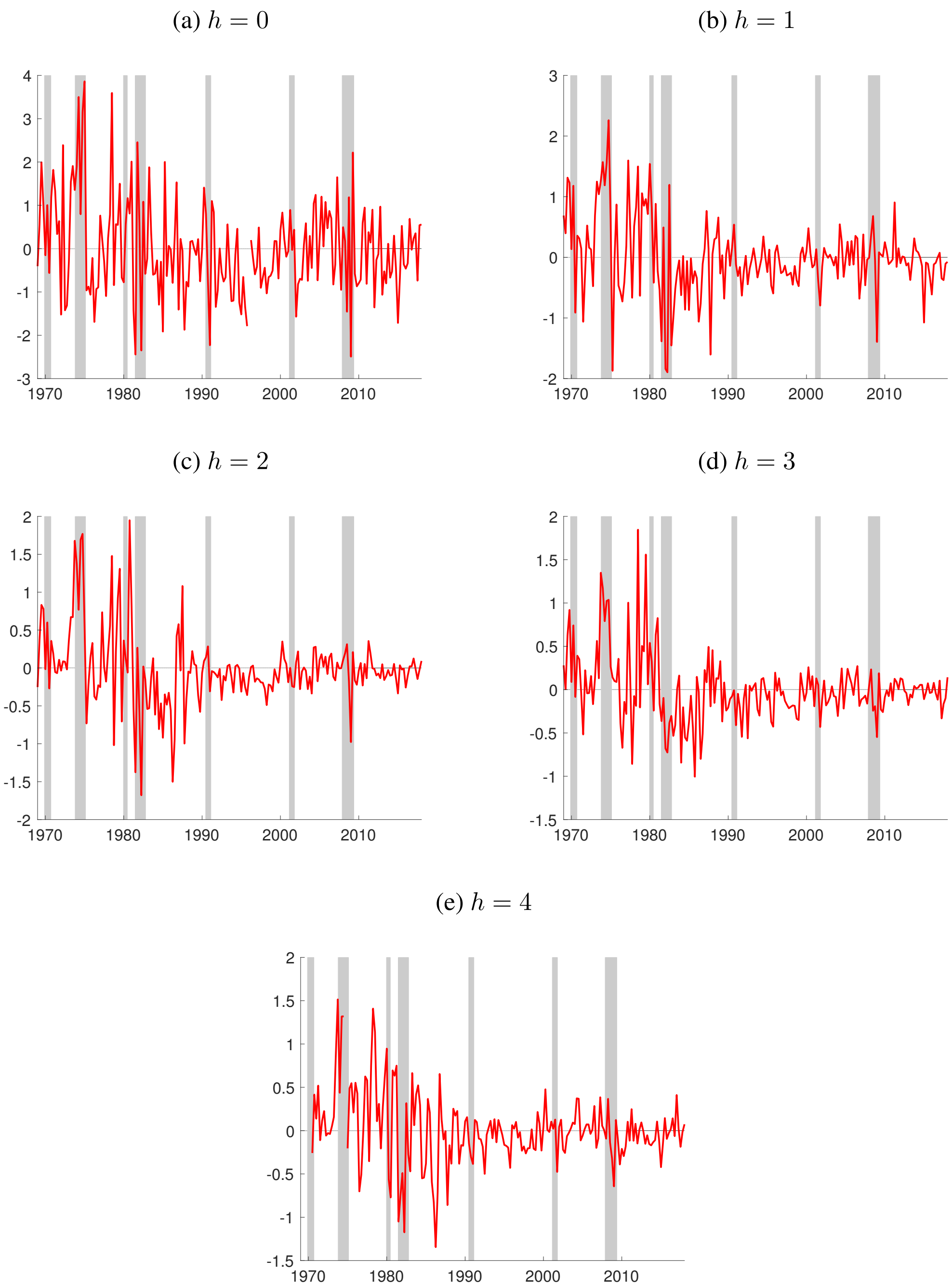

Note: The figure reports the elements of the vector of expectational updates $\boldsymbol{\eta}_t$ used in model estimation. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

Figure A.2: Expectational Updates for the Unemployment Rate




Note: The figure reports the elements of the vector of expectational updates $\boldsymbol{\eta}_t$ used in model estimation. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

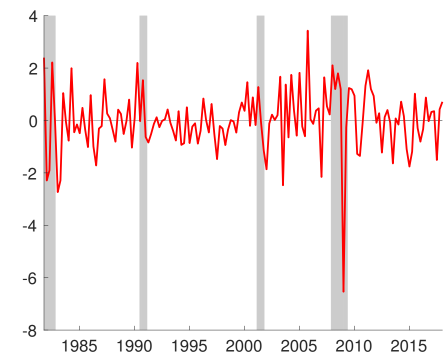
Figure A.3: Expectational Updates for GDP Price Index Inflation



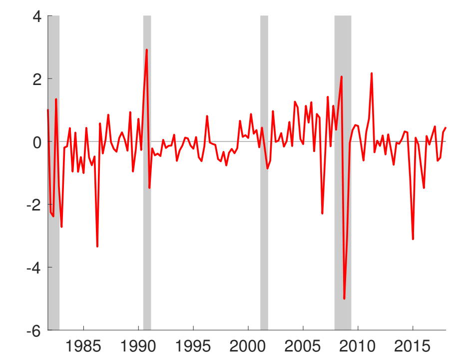

Note: The figure reports the elements of the vector of expectational updates $\boldsymbol{\eta}_t$ used in model estimation. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

Figure A.4: Expectational Updates for CPI Inflation

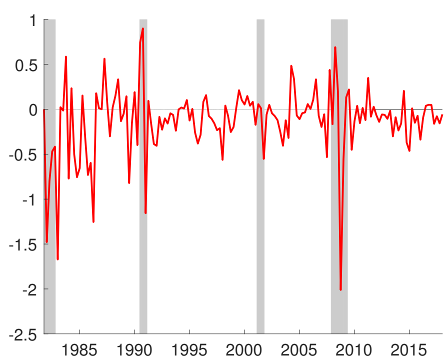
(a) $h = 0$



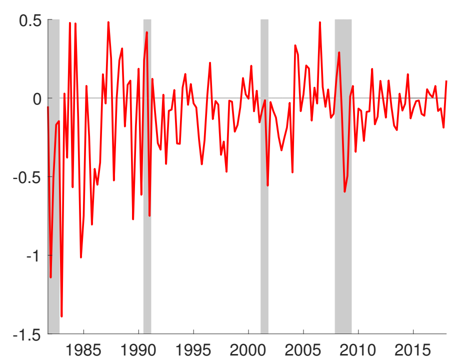

(b) $h = 1$



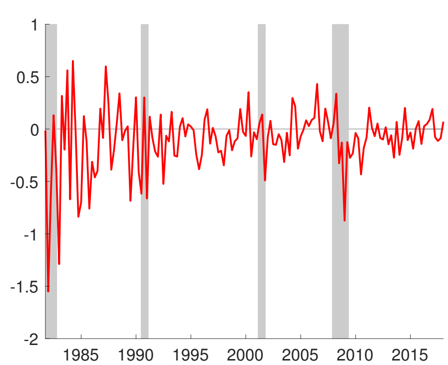

(c) $h = 2$




(d) $h = 3$

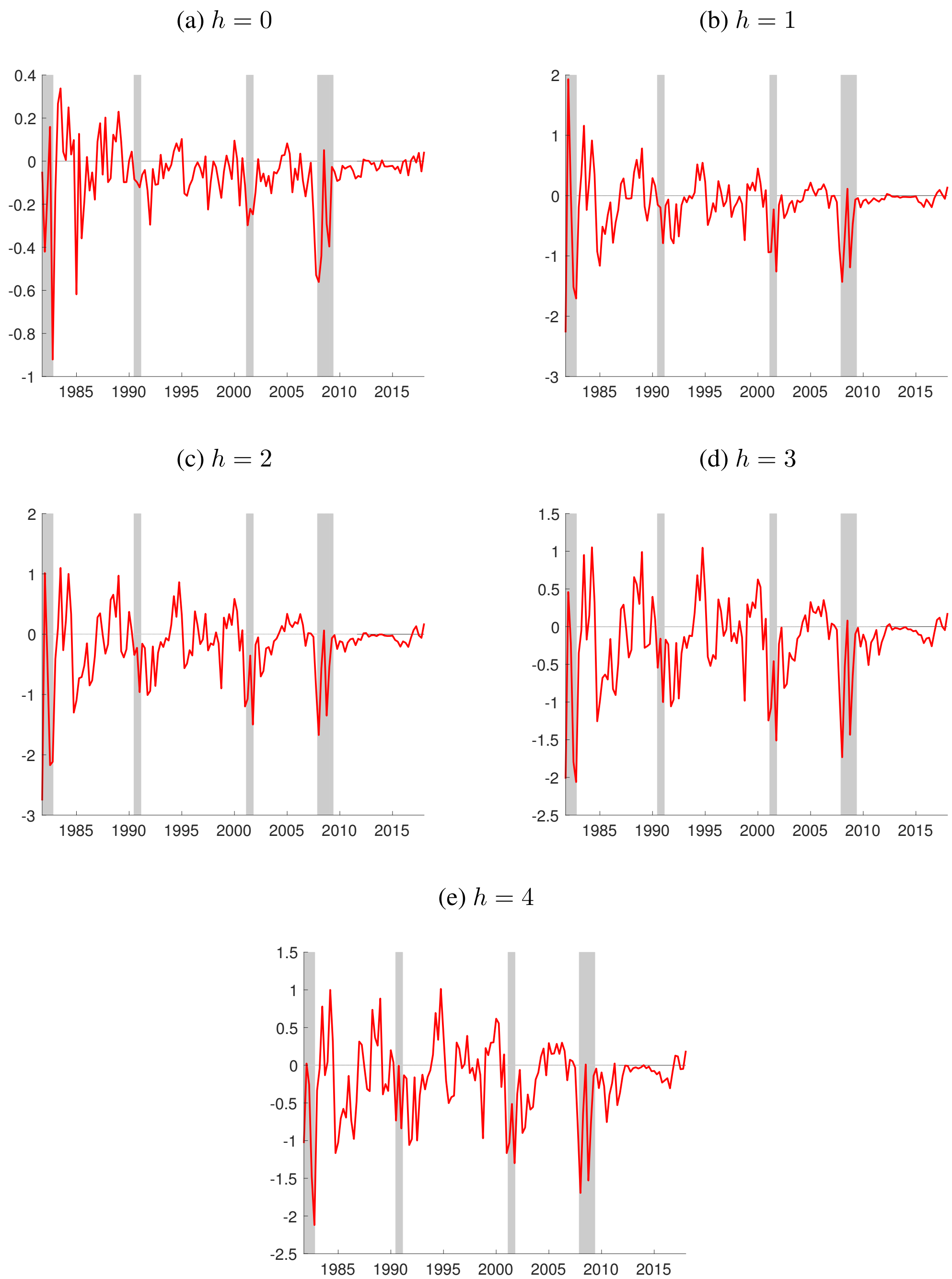



(e) $h = 4$




Note: The figure reports the elements of the vector of expectational updates $\boldsymbol{\eta}_t$ used in model estimation. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

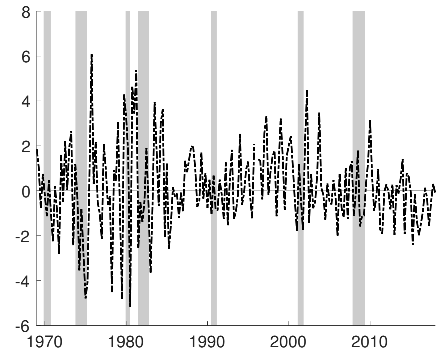
Figure A.5: Expectational Updates for the 3-Month T-bill Rate



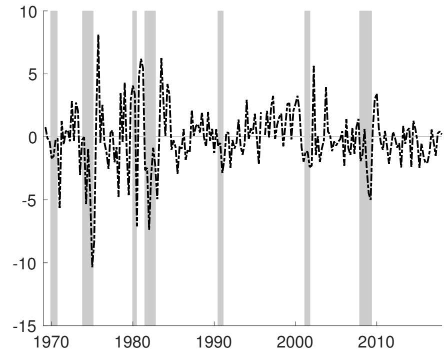

Note: The figure reports the elements of the vector of expectational updates $\boldsymbol{\eta}_t$ used in model estimation. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

Figure A.6: Forecast Errors for Real GDP Growth

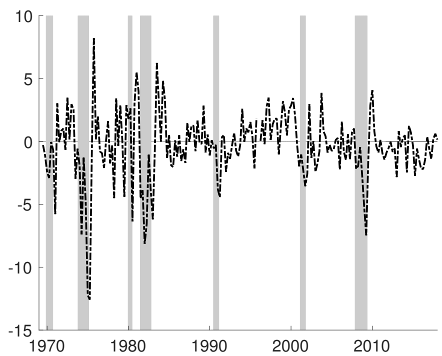
(a) $h = 0$



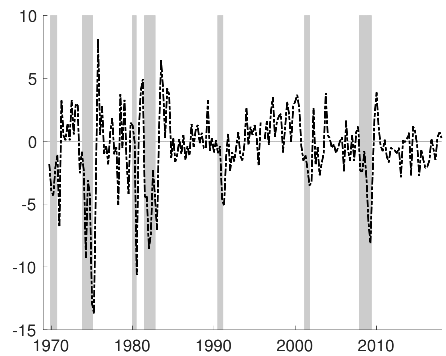

(b) $h = 1$



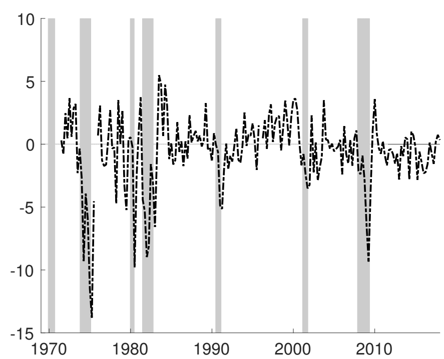

(c) $h = 2$




(d) $h = 3$

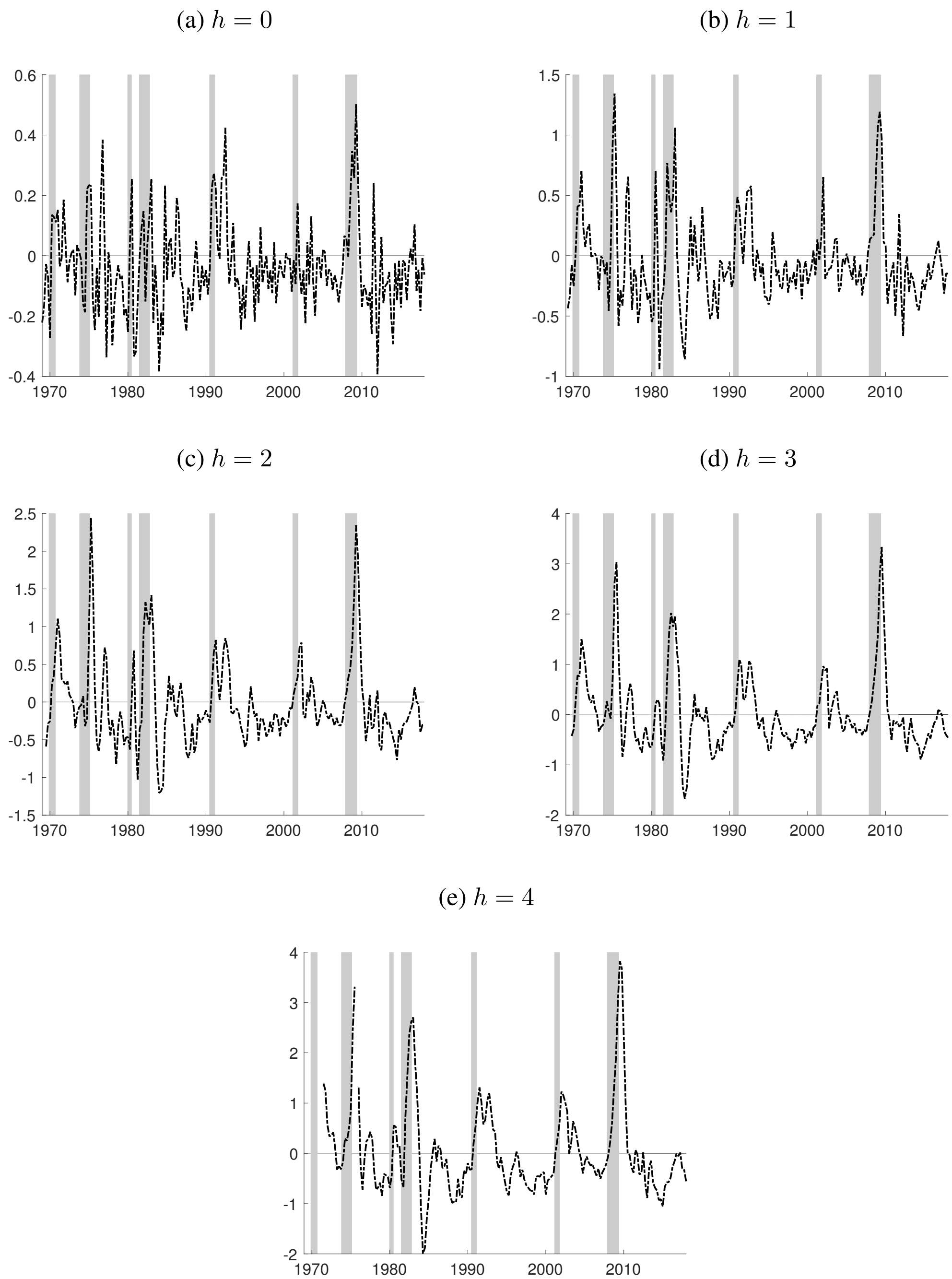



(e) $h = 4$




Note: The figure reports forecast errors $y_{t-1} - E_{t-h-1}y_{t-1} = {}_{t-h-1}e_{t-1}$, plotted at the time when the underlying realized values are observed. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

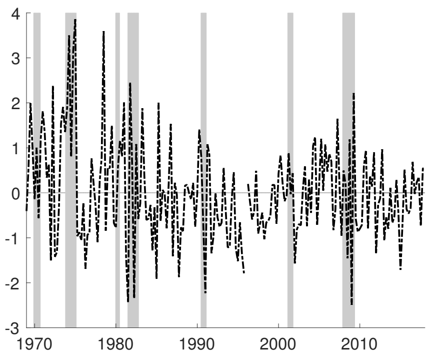
Figure A.7: Forecast Errors for the Unemployment Rate



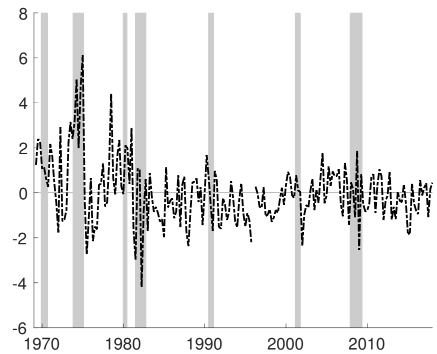

Note: The figure reports forecast errors $y_{t-1} - E_{t-h-1} y_{t-1} = {}_{t-h-1}e_{t-1}$, plotted at the time when the underlying realized values are observed. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

Figure A.8: Forecast Errors for GDP Price Index Inflation

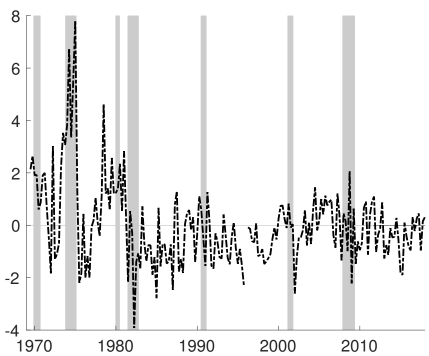
(a) $h = 0$



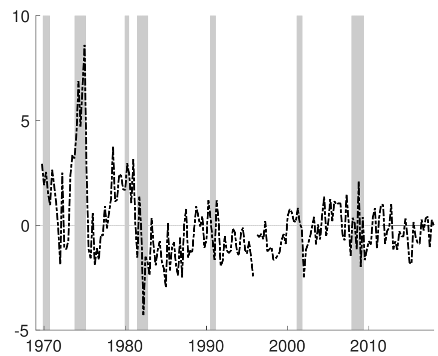

(b) $h = 1$



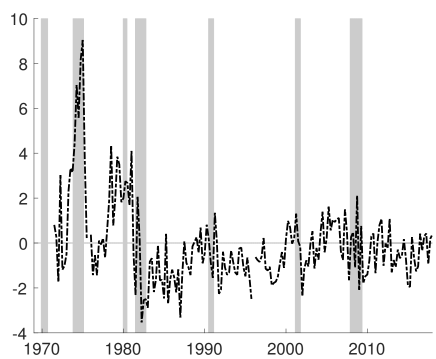

(c) $h = 2$




(d) $h = 3$

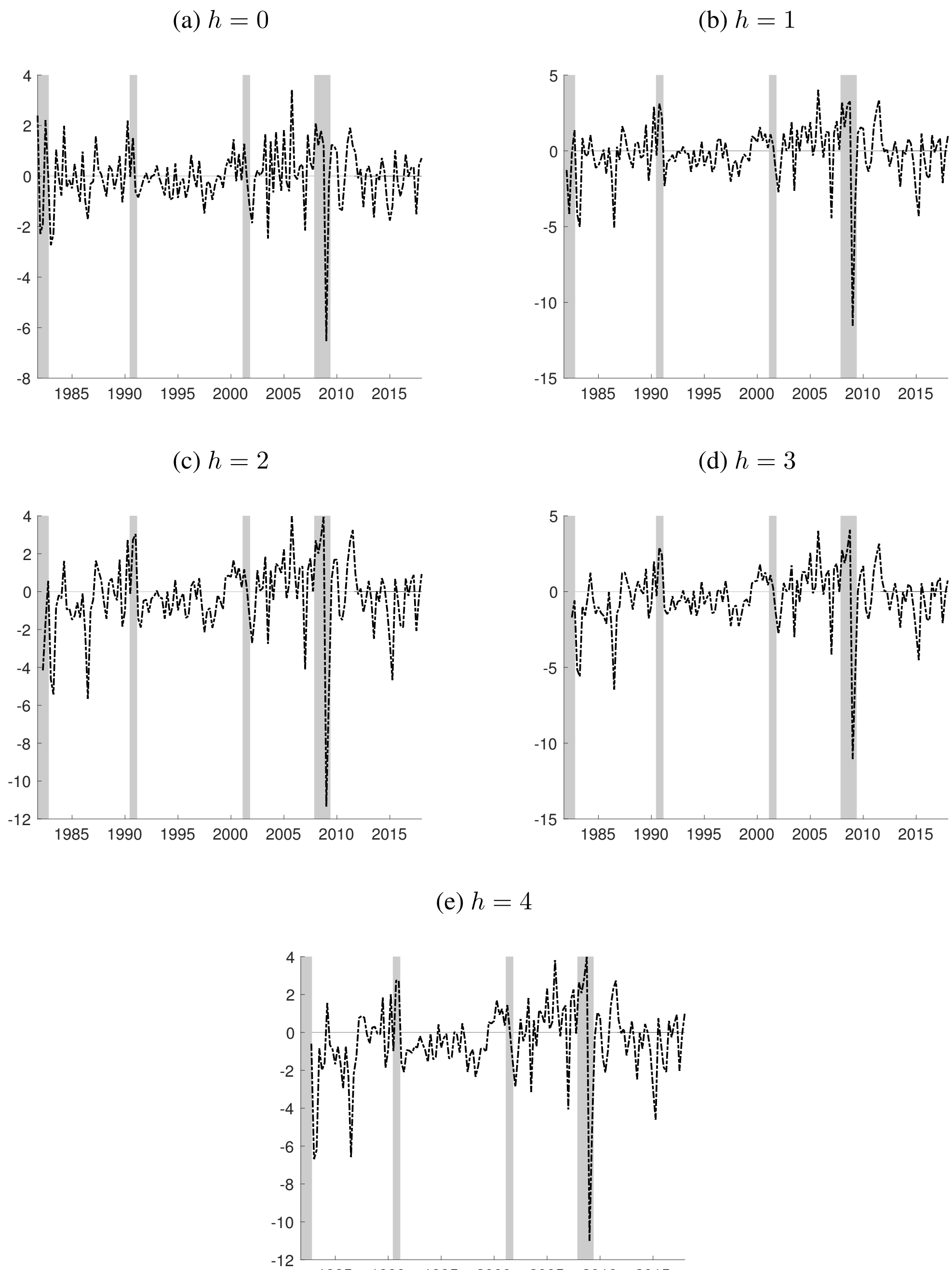



(e) $h = 4$




Note: The figure reports forecast errors $y_{t-1} - E_{t-h-1}y_{t-1} = {}_{t-h-1}e_{t-1}$, plotted at the time when the underlying realized values are observed. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

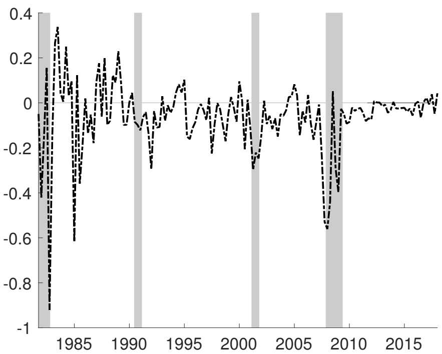
Figure A.9: Forecast Errors for CPI Inflation



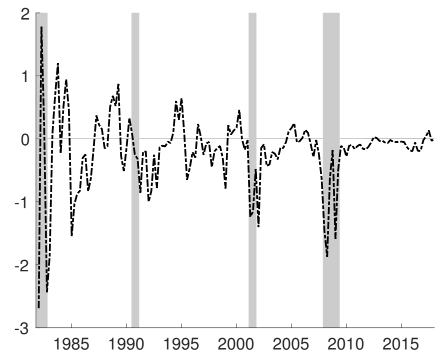

Note: The figure reports forecast errors $y_{t-1} - E_{t-h-1}y_{t-1} = {}_{t-h-1}e_{t-1}$, plotted at the time when the underlying realized values are observed. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

Figure A.10: Forecast Errors for the 3-Month T-bill Rate

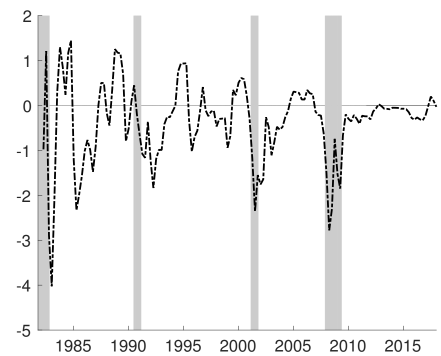
(a) $h = 0$



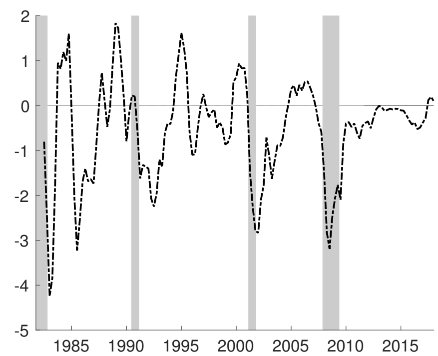

(b) $h = 1$



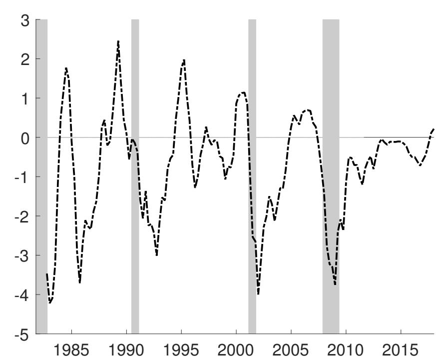

(c) $h = 2$




(d) $h = 3$

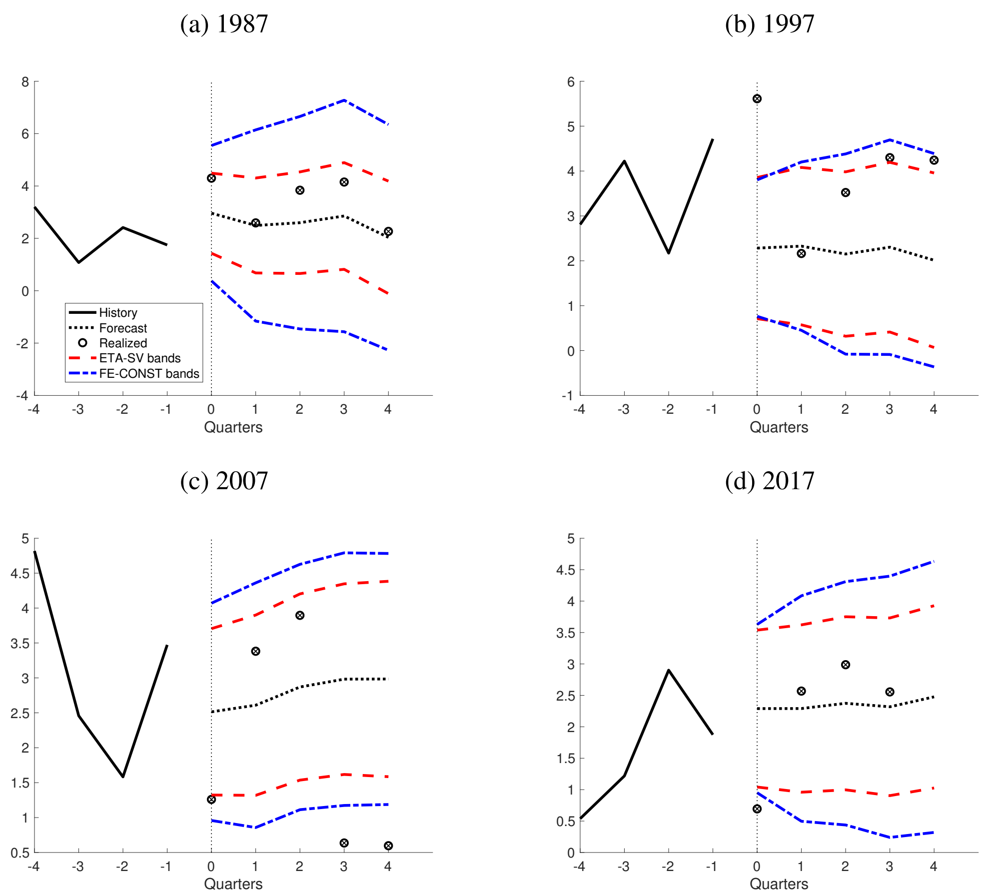



(e) $h = 4$



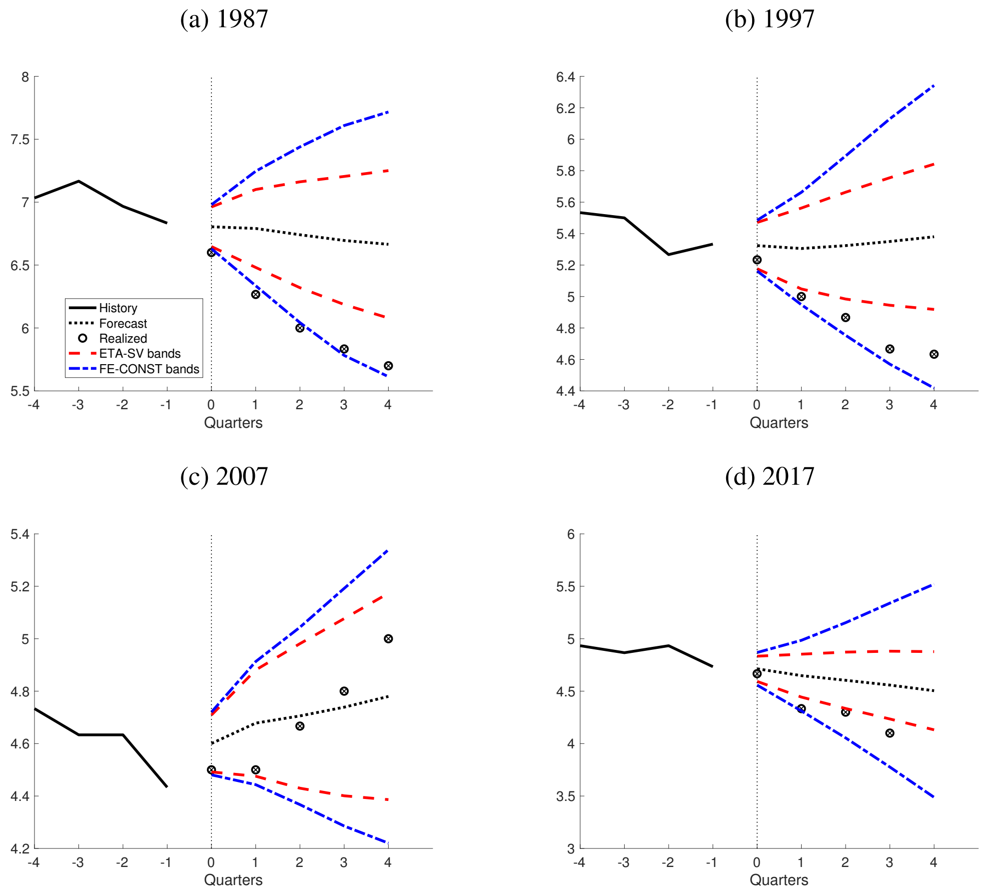

Note: The figure reports forecast errors $y_{t-1} - E_{t-h-1}y_{t-1} = {}_{t-h-1}e_{t-1}$, plotted at the time when the underlying realized values are observed. Expectations reflect the SPF data. NBER recessions are indicated by gray bars.

Figure A.11: Fan Charts for Real GDP Growth



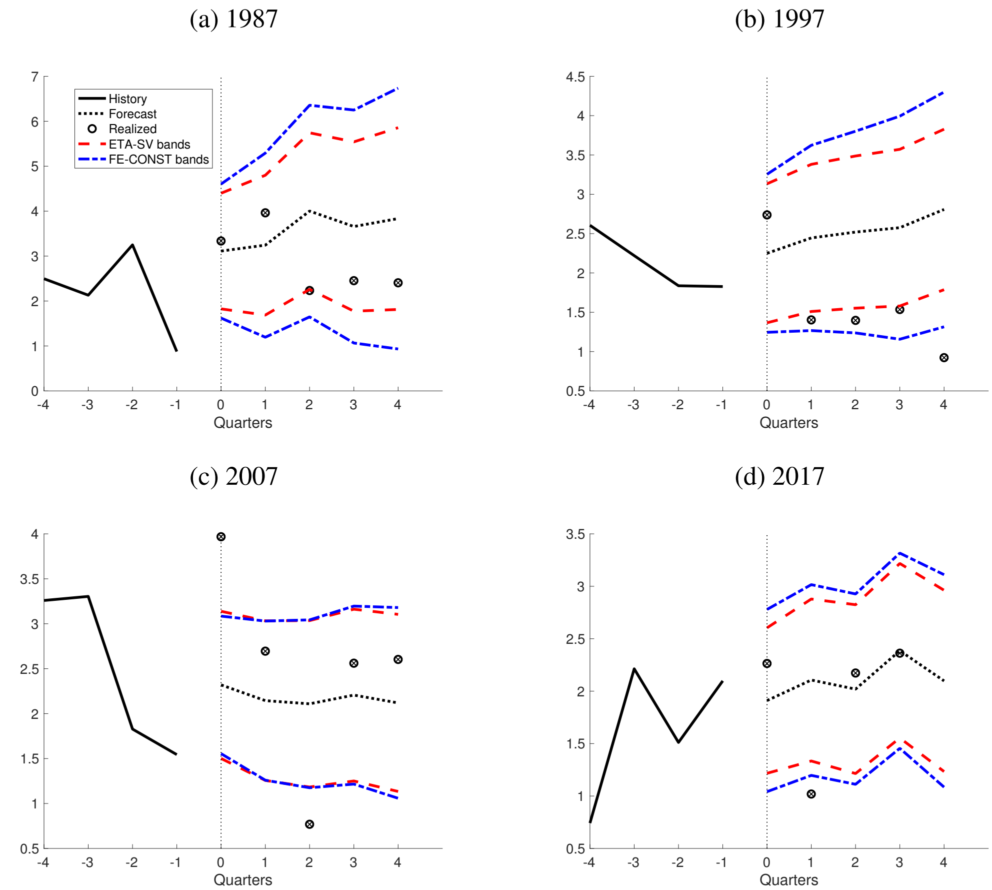

Note: Fan charts generated for the SPF forecasts at selected dates (first quarter of the indicated years). The “fans” reflect one-standard-deviation bands obtained from out-of-sample predictions generated from our baseline ETA-SV model (red dashed lines) and the FE-CONST alternative (blue, dash-dotted lines), respectively.

Figure A.12: Fan Charts for the Unemployment Rate



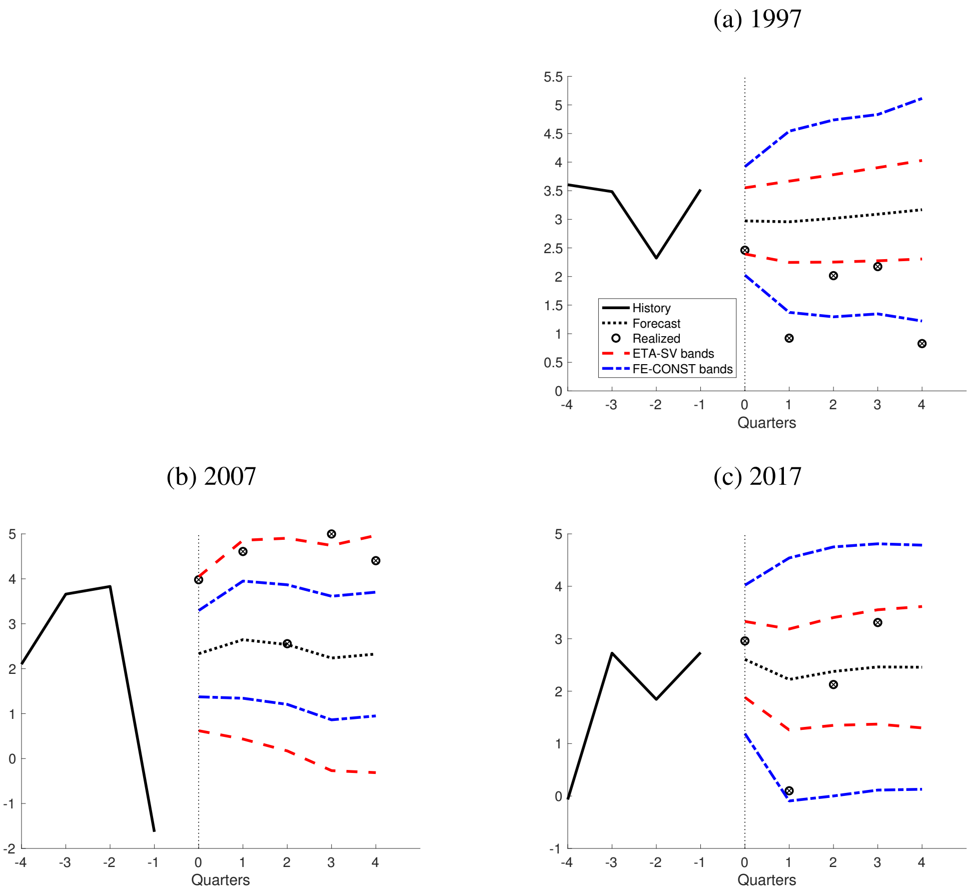

Note: Fan charts generated for the SPF forecasts at selected dates (first quarter of the indicated years). The "fans" reflect one-standard-deviation bands obtained from out-of-sample predictions generated from our baseline ETA-SV model (red dashed lines) and the FE-CONST alternative (blue, dash-dotted lines), respectively.

Figure A.13: Fan Charts for GDP Price Index Inflation




Note: Fan charts generated for the SPF forecasts at selected dates (first quarter of the indicated years). The "fans" reflect one-standard-deviation bands obtained from out-of-sample predictions generated from our baseline ETA-SV model (red dashed lines) and the FE-CONST alternative (blue, dash-dotted lines), respectively.

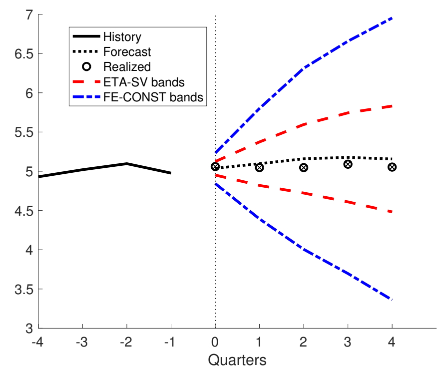
Figure A.14: Fan Charts for CPI Inflation



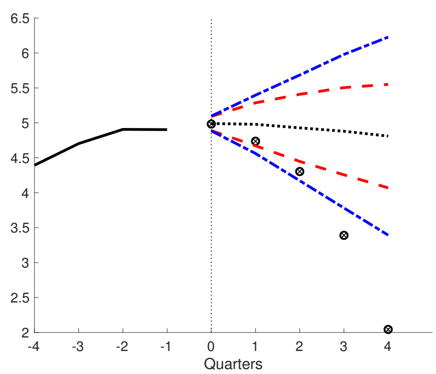

Note: Fan charts generated for the SPF forecasts at selected dates (first quarter of the indicated years). The "fans" reflect one-standard-deviation bands obtained from out-of-sample predictions generated from our baseline ETA-SV model (red dashed lines) and the FE-CONST alternative (blue, dash-dotted lines), respectively.

Figure A.15: Fan Charts for the 3-Month T-bill Rate

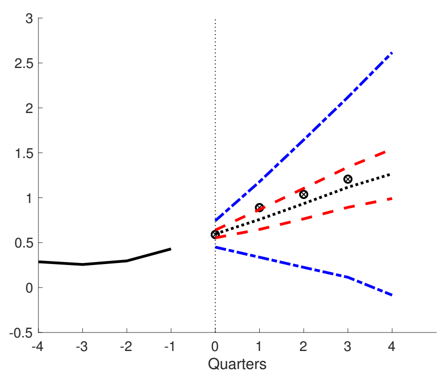
(a) 1997




(b) 2007




(c) 2017




Note: Fan charts generated for the SPF forecasts at selected dates (first quarter of the indicated years). The “fans” reflect one-standard-deviation bands obtained from out-of-sample predictions generated from our baseline ETA-SV model (red dashed lines) and the FE-CONST alternative (blue, dash-dotted lines), respectively.

# IV Additional Results: Greenbook

In this section, we summarize basic results we obtained with forecasts from the Federal Reserve's Greenbook (forecasts from staff at the Board of Governors prepared in advance of FOMC meetings). On balance, our main results obtained with the SPF forecasts are corroborated by estimates with the Greenbook forecasts, although perhaps the efficacy of stochastic volatility is slightly less compelling with the Greenbook than with the SPF forecasts.

The Greenbook forecasts of growth, unemployment, and inflation were obtained from the website of the Federal Reserve Bank of Philadelphia. Although the Federal Reserve prepares forecasts for each FOMC meeting (currently eight meetings per year), we select four forecasts within each year, chosen to align as closely as possible to the timing of the SPF forecast published each quarter. We use forecasts published starting in 1966:Q1 and ending in 2011:Q4 (however, forecasts for CPI inflation do not begin until 1980:Q1). The end of the sample reflects the five year delay in the Federal Reserve's public release of the forecasts. The Greenbook forecasts for the T-bill rate are not provided by the Philadelphia Fed's data files.[3] For comparability, our analysis of the Greenbook forecasts relies on the same choice of horizons as in the case of the SPF. At each forecast origin, we include forecasts spanning five quarterly horizons, from the current quarter through the next four quarters.[4] While the sample of available Greenbook forecasts permits similar start dates to the SPF data, the end date for evaluating is 2011:Q4, reflecting the five-year blackout period for publication.

In the interest of brevity, in examining the robustness of our results to the use of the Greenbook rather than the SPF forecasts, we present the out-of-sample results in tables and omit figures

[3]Studies such as Faust and Wright (2008) and Reifschneider and Tulip (2017) make use of short-term interest rate forecasts from the Greenbook obtained from the Federal Reserve's Board of Governors. However, as discussed in Faust and Wright (2008, 2009), for much of the available history, these forecasts have been tied to conditioning assumptions about monetary policy, rather than unconditional forecasts. Accordingly, we do not include interest rates in our Greenbook assessment.

[4]The maximum forecast horizon available in the Greenbook fluctuates over the course of a calendar year and varies over the course of the data history. Accommodating longer forecast horizons than the $H = 4$ we use would mean accommodating more missing values.

with the full-sample $\boldsymbol{\eta}_t$ and SV estimates. Described in qualitative terms, the full-sample estimates of stochastic volatility with the Greenbook are similar to those for the SPF. In broad terms, along most dimensions, the pattern of interval forecast results for the Greenbook are similar to those for the SPF. First, as in the SPF results, both types of volatility estimates (constant variances with rolling windows and our SV-based estimates) display considerable time variation in the width of the intervals. However, in this dimension, the Greenbook results appear somewhat different in that, up to the mid-2000s, the bands around CPI inflation are fairly stable in width, whereas the SPF-based bands become gradually wider. Second, the width of the confidence bands based on our ETA-SV approach varies more than does the width of intervals based on constant variances. For example, for most variables, the bands widen substantially with the Great Recession and with earlier recessions. Third, across horizons, the contours of the confidence intervals (for a given approach) are very similar.

To quantify the out-of-sample results, Tables A.2 through A.5 reproduce the results generated from the SPF data that are already shown in Tables 1 through 4 of the main paper (coverage rates and CRPS for our baseline ETA-SV model and the VAR-SV variant; henceforth we refer to the latter also as "ETA-VAR-SV"). These tables then also report corresponding results generated from using forecasts obtained from the Federal Reserve's Greenbook.

On balance, coverage rates for one-standard-deviation bands around the Greenbook forecasts based on our stochastic volatility model are comparable to those based on the constant variance approach. For CPI inflation, coverage rates are moderately better with stochastic volatility than in the benchmark. For the unemployment rate, coverage rates also tend to be somewhat closer to the nominal size with the stochastic volatility model than the constant variance approach. But for GDP growth and inflation, coverage rates are quite similar across the two approaches.

The CRPS averages for the Greenbook forecasts show that our ETA-SV model consistently offers some density accuracy gains over the constant variance specification. In broad terms, the gains are comparable to those observed with the SPF forecasts, but a little smaller in most cases.

For example, for GDP growth, the gains range from 3.0 to 9.3 percent with the SPF forecasts and 2.6 to 6.5 percent with the Greenbook forecasts. For the extension to the ETA-VAR-SV models, the patterns in the Greenbook forecasts are similar to those described above for the SPF forecasts. In 68 percent coverage, extending the ETA-SV model to include VAR dynamics does not seem to help much. But in broader density accuracy as captured by the CRPS, the extension to include VAR dynamics reduces accuracy for inflation variables but mostly preserves or extends the gains from ETA-SV for GDP growth and unemployment.

Table A.2: Forecast error coverage rates, including Greenbook results: one-standard-deviation bands

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **Panel A: ETA-SV** | | | | | | |
| RGDP-SPF | 72.06 | 69.63 | 73.13 | 68.42 | 75.76* | 1983:Q4 |
| UNRATE-SPF | 70.80 | 70.59 | 65.93 | 61.19 | 62.41 | 1983:Q4 |
| PGDP-SPF | 73.53 | 71.11 | 71.64 | 70.68 | 71.97 | 1983:Q4 |
| CPI-SPF | 72.09 | 70.59 | 65.48 | 68.67 | 68.29 | 1996:Q3 |
| TBILL-SPF | 76.74* | 77.65* | 70.24 | 63.86 | 50.00** | 1996:Q3 |
| RGDP-Greenbook | 75.00* | 76.42** | 78.69** | 80.99*** | 84.17*** | 1980:Q4 |
| UNRATE-Greenbook | 69.60 | 70.97 | 72.36 | 68.85 | 65.29 | 1980:Q4 |
| PGDP-Greenbook | 70.16 | 77.24** | 77.05** | 78.51** | 76.67 | 1980:Q4 |
| CPI-Greenbook | 68.12 | 69.12 | 68.66 | 68.18 | 72.31 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.09 | 64.71 | 69.05 | 67.47 | 71.95 | 1996:Q3 |
| TBILL-SPF | 79.07* | 88.24*** | 84.52** | 80.72 | 79.27 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.3: Density forecast accuracy as measured by CRPS, including Greenbook results

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **RGDP-SPF** | | | | | | |
| (SV rel.) | 3.01%** | 7.50%*** | 7.96%*** | 9.27%*** | 7.58%*** | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (SV rel.) | 1.78%* | 2.82%** | 3.56%** | 3.44%* | 2.25% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (SV rel.) | 1.03% | 1.41% | 1.83% | 2.59% | 3.00% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (SV rel.) | 1.98% | 2.35% | 1.49% | 1.63% | 2.35% | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (SV rel.) | 11.36%*** | 13.99%*** | 13.00%*** | 9.85%** | 6.86% | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (SV rel.) | 2.59%* | 4.92%*** | 5.62%*** | 6.54%*** | 4.52%** | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (SV rel.) | 1.51% | 2.60%* | 3.78%** | 3.15%** | 1.55% | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (SV rel.) | 1.51% | 1.20% | 2.01% | 2.89% | 3.30%* | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (SV rel.) | 3.27% | 3.08% | 4.57%* | 3.59% | 3.68% | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using SV rather than FE-CONST; positive numbers indicate improvement of SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.4: Forecast error coverage rates, including Greenbook results: one-standard-deviation bands, ETA-VAR-SV specification

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: ETA-VAR-SV** | | | | | | |
| RGDP-SPF | 74.26 | 73.33 | 76.12* | 72.93 | 77.27* | 1983:Q4 |
| UNRATE-SPF | 66.42 | 76.47* | 75.56 | 74.63 | 72.18 | 1983:Q4 |
| PGDP-SPF | 73.53 | 74.07 | 79.85*** | 77.44** | 79.55*** | 1983:Q4 |
| CPI-SPF | 67.44 | 72.94 | 67.86 | 71.08 | 78.05*** | 1996:Q3 |
| TBILL-SPF | 69.77 | 81.18** | 75.00 | 65.06 | 65.85 | 1996:Q3 |
| RGDP-Greenbook | 75.81** | 78.86*** | 82.79*** | 85.12*** | 86.67*** | 1980:Q4 |
| UNRATE-Greenbook | 69.60 | 72.58 | 73.17 | 72.95 | 74.38 | 1980:Q4 |
| PGDP-Greenbook | 70.97 | 78.86*** | 77.87** | 79.34** | 79.17** | 1980:Q4 |
| CPI-Greenbook | 68.12 | 70.59 | 70.15 | 71.21 | 73.85 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.09 | 64.71 | 69.05 | 67.47 | 71.95 | 1996:Q3 |
| TBILL-SPF | 79.07* | 88.24*** | 84.52** | 80.72 | 79.27 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon VAR-SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.5: Density forecast accuracy as measured by CRPS, including Greenbook results, ETA-VAR-SV specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (VAR-SV rel.) | 1.77% | 6.77%*** | 7.36%*** | 7.07%*** | 4.80%** | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (VAR-SV rel.) | 12.15%*** | 11.34%*** | 10.36%*** | 9.02%** | 6.22% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (VAR-SV rel.) | −2.29% | −2.86% | −3.00% | −3.45% | −5.92% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (VAR-SV rel.) | 9.86%** | −0.41% | −1.79% | −2.84% | −3.53% | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (VAR-SV rel.) | 29.74%*** | 25.45%*** | 25.40%*** | 23.76%*** | 21.21%*** | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (VAR-SV rel.) | 0.59% | 4.34%** | 5.40%*** | 5.80%*** | 3.60%* | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (VAR-SV rel.) | 9.13%*** | 8.20%*** | 7.70%** | 6.15%* | 3.10% | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (VAR-SV rel.) | 1.25% | −3.58% | −9.31%** | −11.98%*** | −15.85%** | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (VAR-SV rel.) | 1.68% | 0.63% | 0.74% | −1.80% | −3.23% | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using VAR-SV rather than FE-CONST; positive numbers indicate improvement of VAR-SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

# V Additional Results: Rolling-window choices for FE-CONST

For our baseline SV model, Tables A.6 through A.9 report comparisons of coverage rates and CRPS against the constant variance benchmark — denoted "FE-CONST" since it is estimated directly from observed forecast rather than expectational updates — when the latter is estimated using a rolling window size of 40 and 80 quarters, respectively, rather than the 60-quarter default used elsewhere in this paper. For the sake of comparability, when using the 40-quarter rolling window, the evaluation window used for the comparisons has been left unchanged relative to other results reported in this paper (starting after 60 observations). Consequently, the SV-based results underlying the comparisons reported in Tables A.6 and A.7 are identical to those known from Tables 1 and 2 in the main paper. When using the 80-quarter rolling window, as reported in Tables A.8 and A.9 the evaluation window could begin only as of the 80th observation in our sample.[5]

As indicated in the robustness section of the paper (material repeated here), our baseline results are robust to these modifications of the window choice. Lengthening to 80 observations the rolling window underlying the benchmark constant variance approach does not alter the picture we painted above: the constant variance approach commonly yields coverage rates in excess of the nominal rate of 68 percent. In addition, with the change in the rolling window length, it remains the case that our SV specification offers consistent gains to CRPS accuracy over the constant variance approach. Similarly, shortening the rolling window to 40 observations does not materially change the picture provided by the baseline results, although in forecast coverage, it slightly reduces the advantage of our SV-based model. Density accuracy as measured by the CRPS is only modestly affected by shortening the rolling window from 60 to 40 observations; our SV model-based approach maintains the same consistent advantage described above.

[5] A note on timing conventions in our data set: The data vector $\boldsymbol{\eta}_t$ used in our estimation is observable in real time and involves information about the realization of the lagged value $y_{t-1}$. $T$ observations of $\boldsymbol{\eta}_t$ thus include data of a time zero value of $y_0$ and the underlying data set is covering up to $T+1$ dates.

Table A.6: Forecast error coverage rates: one-standard-deviation bands, FE-CONST using rolling window of 40 observations

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **Panel A: ETA-SV** | | | | | | |
| RGDP-SPF | 72.06 | 69.63 | 73.13 | 68.42 | 75.76* | 1983:Q4 |
| UNRATE-SPF | 70.80 | 70.59 | 65.93 | 61.19 | 62.41 | 1983:Q4 |
| PGDP-SPF | 73.53 | 71.11 | 71.64 | 70.68 | 71.97 | 1983:Q4 |
| CPI-SPF | 72.09 | 70.59 | 65.48 | 68.67 | 68.29 | 1996:Q3 |
| TBILL-SPF | 76.74* | 77.65* | 70.24 | 63.86 | 50.00** | 1996:Q3 |
| RGDP-Greenbook | 75.00* | 76.42** | 78.69** | 80.99*** | 84.17*** | 1980:Q4 |
| UNRATE-Greenbook | 69.60 | 70.97 | 72.36 | 68.85 | 65.29 | 1980:Q4 |
| PGDP-Greenbook | 70.16 | 77.24** | 77.05** | 78.51** | 76.67 | 1980:Q4 |
| CPI-Greenbook | 68.12 | 69.12 | 68.66 | 68.18 | 72.31 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 72.79 | 73.33 | 73.13 | 75.19 | 75.00 | 1983:Q4 |
| UNRATE-SPF | 71.53 | 77.94** | 81.48*** | 82.84*** | 84.21*** | 1983:Q4 |
| PGDP-SPF | 73.53 | 74.81* | 70.90 | 72.93 | 76.52* | 1983:Q4 |
| CPI-SPF | 73.26 | 67.06 | 65.48 | 67.47 | 68.29 | 1996:Q3 |
| TBILL-SPF | 79.07* | 85.88*** | 84.52*** | 81.93* | 78.05 | 1996:Q3 |
| RGDP-Greenbook | 75.00* | 77.24** | 73.77 | 71.90 | 70.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 69.35 | 69.92 | 71.31 | 76.03 | 1980:Q4 |
| PGDP-Greenbook | 75.00 | 73.98 | 70.49 | 75.21 | 71.67 | 1980:Q4 |
| CPI-Greenbook | 63.77 | 63.24 | 58.21 | 54.55* | 49.23** | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 40 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.7: Density forecast accuracy as measured by CRPS, FE-CONST using rolling window of 40 observations

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (SV rel.) | 2.42%** | 6.03%*** | 5.81%*** | 6.67%*** | 5.50%*** | 1983:Q4 |
| (FE-CONST) | 0.81 | 1.00 | 1.07 | 1.13 | 1.14 | |
| **UNRATE-SPF** | | | | | | |
| (SV rel.) | 1.71%* | 2.53%* | 3.44%* | 3.53% | 2.59% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.44 | |
| **PGDP-SPF** | | | | | | |
| (SV rel.) | 0.78% | 0.07% | −0.08% | 0.26% | 0.36% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.55 | 0.58 | 0.62 | 0.66 | |
| **CPI-SPF** | | | | | | |
| (SV rel.) | 2.47% | 3.42% | 2.82% | 2.87% | 3.21% | 1996:Q3 |
| (FE-CONST) | 0.67 | 1.06 | 1.11 | 1.11 | 1.11 | |
| **TBILL-SPF** | | | | | | |
| (SV rel.) | 11.72%*** | 12.78%*** | 11.23%*** | 8.16%** | 5.42% | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.39 | 0.57 | 0.75 | |
| **RGDP-GB** | | | | | | |
| (SV rel.) | 2.15%* | 3.46%** | 3.91%** | 4.68%** | 2.77% | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.20 | 1.30 | 1.37 | 1.34 | |
| **UNRATE-GB** | | | | | | |
| (SV rel.) | 1.64% | 3.37%** | 4.41%*** | 3.40%** | 1.71% | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.41 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (SV rel.) | 1.18% | 0.13% | −0.34% | −0.12% | 0.42% | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.60 | 0.63 | 0.67 | |
| **CPI-GB** | | | | | | |
| (SV rel.) | 0.66% | 2.09% | 3.56% | 2.77% | 2.77% | 1994:Q4 |
| (FE-CONST) | 0.53 | 1.09 | 1.21 | 1.23 | 1.21 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using SV rather than FE-CONST; positive numbers indicate improvement of SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 40 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.8: Forecast error coverage rates: one-standard-deviation bands, FE-CONST using rolling window of 80 observations

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: ETA-SV** | | | | | | |
| RGDP-SPF | 69.83 | 66.96 | 69.30 | 62.83 | 69.64 | 1988:Q4 |
| UNRATE-SPF | 70.09 | 71.55 | 65.22 | 60.53 | 61.06 | 1988:Q4 |
| PGDP-SPF | 74.14 | 69.57 | 70.18 | 69.03 | 70.54 | 1988:Q4 |
| CPI-SPF | 72.73 | 73.85 | 71.88 | 74.60 | 75.81 | 2001:Q3 |
| TBILL-SPF | 78.79** | 80.00* | 76.56 | 69.84 | 56.45 | 2001:Q3 |
| RGDP-Greenbook | 74.04 | 77.67*** | 78.43** | 79.21** | 80.00** | 1985:Q4 |
| UNRATE-Greenbook | 67.62 | 70.19 | 72.82 | 67.65 | 64.36 | 1985:Q4 |
| PGDP-Greenbook | 70.19 | 74.76 | 75.49 | 76.24* | 74.00 | 1985:Q4 |
| CPI-Greenbook | 63.27 | 62.50 | 61.70 | 63.04 | 71.11 | 1999:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 81.03*** | 80.00*** | 78.07** | 79.65** | 79.46** | 1988:Q4 |
| UNRATE-SPF | 74.36 | 82.76*** | 85.22*** | 85.09*** | 84.07*** | 1988:Q4 |
| PGDP-SPF | 80.17*** | 82.61*** | 82.46*** | 83.19*** | 83.04*** | 1988:Q4 |
| CPI-SPF | 65.15 | 66.15 | 71.88 | 69.84 | 67.74 | 2001:Q3 |
| TBILL-SPF | 84.85*** | 90.77*** | 89.06*** | 85.71** | 83.87* | 2001:Q3 |
| RGDP-Greenbook | 82.69*** | 82.52*** | 79.41** | 78.22* | 81.00** | 1985:Q4 |
| UNRATE-Greenbook | 68.57 | 75.00 | 82.52** | 82.35** | 83.17** | 1985:Q4 |
| PGDP-Greenbook | 80.77*** | 78.64** | 78.43* | 76.24 | 76.00 | 1985:Q4 |
| CPI-Greenbook | 61.22 | 54.17* | 55.32 | 50.00* | 51.11* | 1999:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 80 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.9: Density forecast accuracy as measured by CRPS, FE-CONST using rolling window of 80 observations

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (SV rel.) | 2.00% | 5.79%*** | 5.58%** | 5.63%** | 4.37%* | 1988:Q4 |
| (FE-CONST) | 0.79 | 0.99 | 1.08 | 1.16 | 1.18 | |
| **UNRATE-SPF** | | | | | | |
| (SV rel.) | 1.63%* | 3.66%*** | 3.57%** | 2.60% | 1.06% | 1988:Q4 |
| (FE-CONST) | 0.08 | 0.16 | 0.25 | 0.34 | 0.44 | |
| **PGDP-SPF** | | | | | | |
| (SV rel.) | 1.28% | 1.84% | 3.19%* | 3.52%* | 3.51% | 1988:Q4 |
| (FE-CONST) | 0.49 | 0.54 | 0.56 | 0.60 | 0.65 | |
| **CPI-SPF** | | | | | | |
| (SV rel.) | 2.03% | 1.50% | 1.27% | 1.98% | 2.45% | 2001:Q3 |
| (FE-CONST) | 0.74 | 1.18 | 1.20 | 1.19 | 1.19 | |
| **TBILL-SPF** | | | | | | |
| (SV rel.) | 12.47%*** | 15.47%*** | 15.11%*** | 11.98%** | 8.12% | 2001:Q3 |
| (FE-CONST) | 0.07 | 0.22 | 0.36 | 0.52 | 0.69 | |
| **RGDP-GB** | | | | | | |
| (SV rel.) | 2.30% | 5.72%*** | 6.19%*** | 6.45%** | 4.13%* | 1985:Q4 |
| (FE-CONST) | 0.82 | 1.09 | 1.21 | 1.30 | 1.28 | |
| **UNRATE-GB** | | | | | | |
| (SV rel.) | 2.36% | 3.68%** | 4.05%** | 2.28% | 0.56% | 1985:Q4 |
| (FE-CONST) | 0.10 | 0.19 | 0.27 | 0.37 | 0.46 | |
| **PGDP-GB** | | | | | | |
| (SV rel.) | 2.36% | 2.46%* | 4.33%* | 5.11%* | 5.91%** | 1985:Q4 |
| (FE-CONST) | 0.48 | 0.58 | 0.59 | 0.62 | 0.65 | |
| **CPI-GB** | | | | | | |
| (SV rel.) | 1.47% | 0.67% | 2.79% | 2.66% | 3.01% | 1999:Q4 |
| (FE-CONST) | 0.66 | 1.35 | 1.48 | 1.47 | 1.45 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using SV rather than FE-CONST; positive numbers indicate improvement of SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 80 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

# VI Other Model Specifications

## VI.1 AR(1) specification for stochastic volatilities

Our baseline stochastic volatility (SV) model assumes that log-variances evolve as random walks; see equation (10) of the main paper. Among others, our choice is grounded in the results of Clark and Ravazzolo (2015); in a comparison of forecasting models using different specifications of heteroscedasticity, these authors found the random-walk specification for SV to perform generally better than others. Nevertheless, as a robustness check, we have also re-estimated our results from a version of the SV model where log-variances evolve as stationary AR(1) processes rather than random walks.

For this model variant, equation (10) of the main paper is replaced by the following:

$$\log(\lambda_{i,t}) = (1-\rho_i)\ \log(\bar{\lambda}_i) + \rho_i\ \log(\lambda_{i,t-1}) + \nu_{i,t}, \quad i = 1, \ldots, H+1, \quad |\rho_i| < 1. \tag{A.15}$$

The average levels of the log-variances, $\bar{\lambda}_i$, are estimated as part of the KSC state space, by including $\bar{\lambda}_i$ as additional state variables (in the form of unit-root processes without innovations). The priors for $\bar{\lambda}_i$ are identical to those used for the initial levels $\bar{\lambda}_{i,0}$ in the random-walk case, $\bar{\lambda}_i \sim N(\log(0.5^2), 10)$. For the AR(1) persistence coefficient, we use the same normal prior as Clark and Ravazzolo (2015): $\rho_i \sim N(0.8, 0.2^2)$.

Similar to the random-walk case, the AR(1) innovations are assumed to be correlated across horizons $i$ (using the same inverse-Wishart prior as in the baseline SV case). Consequently, the system of AR(1) equations given by (A.15) across different $i$ constitutes a system of seemingly unrelated regressions (SUR) and its estimation is embedded as an additional Gibbs step in our MCMC sampler using the conjugate prior-posterior relationships described, for example, by Koop (2003). We apply rejection sampling and accept draws for $\rho_i$ that are inside the unit circle.

Results obtained from comparing this model version against the constant variance benchmark are reported in Tables A.10 and A.11. The results (in terms of coverage rates and density forecast accuracy as measured by CRPS) are similar to our baseline specification; in comparison to the constant variance benchmark, neither specification strictly dominates the other across all variables and at every horizon.

## VI.2 Factor model for stochastic volatilities

Our baseline SV estimates, shown in Figures 1 and 2 of the main paper, display a fair amount of commonality. In this baseline model, that commonality is captured by the estimated correlations between log-variance innovations $\nu_{i,t}$ across $i$. As a further robustness check, we considered an alternative version of our basic SV model where variations in SV across different elements of the vector $\tilde{\boldsymbol{\eta}}_t$ are solely driven by one common factor.

Replacing the random-walk specification for log-variances provided by equation (10) in the main paper, we assume the following single-factor SV structure:[6]

$$\log \tilde{\eta}_{i,t}^2 = \log \lambda_{i,t} + \log \varepsilon_{i,t}^2 \qquad i = 1, \dots, H+1, \tag{9}$$

$$\log(\lambda_{i,t}) = \log(\bar{\lambda}_i) + \beta_i f_t \tag{A.16}$$

$$f_t = f_{t-1} + \nu_{f,t} \qquad \nu_{f,t} \sim N(0, \phi_f) \tag{A.17}$$

For given values of the factor slopes $\beta_i$ and factor innovation variance $\phi_f$, this system can be estimated with a slightly modified version of the (approximate) state space described by Kim, Shephard, and Chib (1998) [henceforth, KSC] with an unchanged measurement equation (9), reproduced above, and modified state equations.[7] As a matter of normalization, we set $f_0 = 0$ and $\beta_1 = 1$.

[6]This single-factor structure is a simplified version of the SV factor model described by Carrierio, Clark, and Marcellino (2015).

[7]Constants, $\log(\bar{\lambda}_i)$, are tracked as part of the state space in the form of unit root processes without innovations. In addition, we use the 10-state mixture of normals approximation presented by Omori, et al. (2007) rather than the original seven states known from KSC.

Analogously to our baseline SV model, the prior for $\phi_f$ is a (univariate) inverse Wishart with 3 degrees of freedom and centered around a mean of $0.2^2$. The prior for the SV constants is $\log(\bar{\lambda}_i) \sim N(\log(0.5^2), 10)$ and the slopes have normal priors $\beta_i \sim N(1, 0.5)$ as well.[8]

Results obtained from comparing this model version against the constant variance benchmark are reported in Tables A.12 and A.13. The results (in terms of coverage rates and density forecast accuracy as measured by CRPS) are similar to our baseline specification; in comparison to the constant variance benchmark, neither specification strictly dominates the other across all variables and at every horizon.

## VI.3 Joint model of expectational updates for multiple variables

Throughout this paper, we have focused on characterizing uncertainty around forecasts for each variable separately. A wide body of literature suggests that there is also substantial commonality in time variation of forecast uncertainty across different macroeconomic variables; see, for example, Jurado, Ludvigson, and Ng (2015), and Carriero, Clark, and Marcellino (2016). Indeed, the SV estimates shown in Figures 1 and 2 display notable commonality not only across forecast horizons but also across variables.

To assess the usefulness of pooling information across variables in the estimation, we have applied the baseline SV model as well as its SV-Factor variant to a joint data set involving the expectational updates from the SPF for the unemployment rate, real GDP growth, and GDP price index inflation. (Forecast data for CPI inflation and the 3-month T-bill were omitted because of the more limited availability of historical forecast data for those variables.) The resulting model variants will be denoted JOINT-SV and JOINT-SV-Factor, respectively.

Except for the use of an augmented data vector, $\boldsymbol{\eta}_t$, and different lag-length choices, the specification of the JOINT models is identical to their single-variable counterparts described

[8] Given the conjugate priors, the slopes $\beta_i$ and factor innovation variance $\phi_f$ are straightforward to estimate via additional Gibbs steps in our MCMC sampler.

above and in the main paper. In the single-variable case, $\boldsymbol{\eta}_t$ consists of $H + 1$ expectational updates for forecasts of one variable.[9] In the JOINT case, $\boldsymbol{\eta}_t$ now contains all $(H + 1) \cdot N_y$ expectational updates in forecasts for $N_y$ variables at $H + 1$ different horizons. In the JOINT-SV model, cross-variable (and cross-horizon) commonality between SV processes is captured via the estimated variance covariance matrix $\boldsymbol{\Phi}$ of the log-variance innovations $\nu_{i,t}$ (with $i = 1, 2, \ldots, N_y \cdot H$). In the JOINT-SV-Factor model, commonality in SV process is again captured by a single factor. In both cases, correlation between the *levels* of $\boldsymbol{\eta}_t$ is captured via the lower-triangular matrix described in equation (7) of the main paper. In our specific application, using only forecasts for the unemployment rate, real GDP Growth, and GDP price index inflation, we have $N_y = 3$, $H = 4$. Observations are stacked by horizon and then by variable using the aforementioned order of variables.

Results comparing this model version against the constant variance benchmark are reported in Tables A.14 and A.15 — assuming separate, but correlated SV processes for each element of $\boldsymbol{\eta}_t$ — and Tables A.16 and A.17 — assuming a single common SV factor drives time variation in the volatilities of each expectational update across variables and horizons. As summarized in the paper's robustness section, coverage and CRPS results from these trivariate specifications are similar to those from our baseline analysis. In coverage, the trivariate specifications perform a little worse than our baseline, and in CRPS, they are a little better in some cases and worse in others.

In the context of the JOINT models, it might be worthwhile to note some important aspects in which our work is distinct from some of the literature on measuring uncertainty and its macroeconomic effects. For example, Jurado, Ludvigson, and Ng (2015) use factor-augmented autoregressive models to capture the conditional means of macro variables and obtain estimates of stochastic volatility, abstracting from considerations of real-time data. Taking the resulting volatility estimates as given, they go on to define uncertainty as an average across variables of

[9]Please recall that there are forecasts for horizons $h = 0, 1, \ldots, H$ (including the nowcast).

(ex post) forecast error variances and assess its macroeconomic effects with a vector autoregression. We instead take point forecasts as given from a source such as the SPF — remaining agnostic about the data-generating process of the underlying data as well as details of the forecasting model — and focus on the measurement of possibly time-varying uncertainty around each forecast, in a real-time, ex ante data setting.[10] With alternative uncertainty estimates in hand, we evaluate their efficacy.

## VI.4 Constant variance estimates derived from expectational updates

Estimates from the constant variance benchmark used in model comparisons throughout this paper are directly generated from observed forecast error data by computing root-mean square errors, without transformations into expectational updates and without the martingale-difference sequence (MDS) assumption embedded in our baseline SV model. To better parse out the role of the MDS assumption relative to the role of SV, we have also estimated a constant variance version of the ETA-SV model, also written in terms of expectational updates $\boldsymbol{\eta}_t$. Estimation of this "ETA-CONST" model uses an MCMC sample similar to the ETA-SV case, but assuming a time-invariant distribution of expectational updates, $\boldsymbol{\eta}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma})$. We employ Bayesian methods to estimate this model within the real-time setup described in the main paper, assuming a diffuse inverse-Wishart prior for $\boldsymbol{\Sigma}$.[11] As in the FE-CONST case, ETA-CONST is estimated using 60-quarter rolling windows of data (rather than the growing data windows used for the SV models).

Results obtained from comparing this model version against the constant variance benchmark that uses data on forecast errors directly ("FE-CONST") are reported in Tables A.18 and A.19.

[10] Jo and Sekkel (2017) also take the SPF forecasts as given and obtain a measure of macroeconomic uncertainty from a factor model with stochastic volatility applied to the one-step-ahead forecast errors of a few SPF variables. They use the estimate to assess its macroeconomic effects rather than to assess the accuracy of uncertainty estimates.

[11] Draws for $\boldsymbol{\Sigma}$ are generated via a Gibbs step that employs the usual conjugate prior-posterior relationships; this Gibbs step replaces drawing coefficients of the Cholesky factor $\boldsymbol{A}$ in equation (7) as well as the KSC sampling steps in our MCMC sampler.

Exploiting the martingale-difference assumption embedded into the ETA-CONST approach does indeed yield notable improvements in coverage rates and density forecast accuracy, at least for a number of variables. However, as noted already in Section 5 of the main paper, the measured improvements in coverage rates and accuracy of density forecasts relative to FE-CONST tend to be consistently smaller than those reported for the baseline SV model as reported in the paper's Tables 1 and 2 and repeated in Tables A.2 and A.3. As reported in Table A.20, directly comparing the CRPS results for our preferred ETA-SV specification against the alternative ETA-CONST baseline shows that our preferred model yields more accurate density forecasts, often significantly so. But the gains are more modest than when our preferred model is compared against the FE-CONST baseline as in the paper. As we note in the paper's robustness section, in an overall sense, our methodological innovation has two components, one of which is the use of the expectational updates and the other is the use of stochastic volatility (with the former enabling the latter), and both components appear helpful for the problem at hand.

## VI.5 VAR with stochastic volatilities applied directly to forecast errors

Our preferred ETA-SV model differs from the constant variance benchmark (FE-CONST) in two aspects: Time-varying uncertainty modeled via stochastic volatilities and the MDS assumption embedded in the setup written in terms of expectational updates. Provided the MDS assumption is true (or provides a decent approximation to the data), the setup in expectational updates optimally removes serial correlation between observed forecast errors due to overlapping forecast horizons. As a further robustness check, we have also applied a VAR system with stochastic volatility in its innovations directly to data on forecast errors. This model variant will henceforth be referred to as "FE-VAR(p)-SV."

Apart from input data and lag length, the specification of the VAR and its estimation is identical to the generalized model without MDS assumption described in Section 3.3 of the main paper, which was applied to the vector of expectational updates $\boldsymbol{\eta}_t$.[12] The FE-VAR(p)-SV model is applied to the vector of forecast errors $\boldsymbol{e}_t$ defined in equation (6) of the main paper.[13] In light of the overlapping forecast windows, observations of $\boldsymbol{e}_t$ should have stronger serial correlation than data on expectational updates $\boldsymbol{\eta}_t$ and over longer lags.[14] Indeed, preliminary analysis of information criteria computed for (homoscedastic) VARs in $\boldsymbol{e}_t$ with different lag lengths suggest values of 2 but also up to $4$ or $5$ lags. Below we report results for FE-VAR(p)-SV estimates using lag-length choices of $p = 2$ as well as $p = 5$.

Results obtained from comparing the constant variance benchmark, FE-CONST, against FE-VAR(2)-SV are reported in Tables A.21 and A.22 while Tables A.23 and A.24 provide results based on $p = 5$. Indeed, compared to the constant variance benchmark, the FE-VAR(p)-SV model fares somewhat better in terms of coverage rates and forecast density accuracy (though not uniformly). But using either metric, the performance of FE-VAR(p)-SV is typically inferior to ETA-SV or ETA-VAR-SV models. In most cases, coverage rates are higher (less accurate) with the FE-VAR(p)-SV model than the baseline SV model; by the CRPS measure, the FE-VAR(p)-SV model is less accurate than the baseline SV model for most, although not all, variables.

[12]As noted already in the main paper, this VAR-SV model has the same form as the model of Cogley and Sargent (2005) that has by now been considered in a number of forecasting studies.

[13]Note that real-time readings for realized values $y_t$ are available only in quarter $t+1$, and our real-time analysis has been adjusted accordingly.

[14]Under the MDS assumption, there should be no serial correlation in $\boldsymbol{\eta}_t$. In contrast, forecast error windows are overlapping for up to $H-1$ lags.

Table A.10: Forecast error coverage rates, SV-AR(1) specification: one-standard-deviation bands

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: SV-AR(1)** | | | | | | |
| RGDP-SPF | 72.06 | 71.11 | 75.37 | 77.44* | 78.03** | 1983:Q4 |
| UNRATE-SPF | 71.53 | 70.59 | 69.63 | 62.69 | 63.91 | 1983:Q4 |
| PGDP-SPF | 80.15*** | 80.00*** | 82.84*** | 81.20*** | 79.55*** | 1983:Q4 |
| CPI-SPF | 68.60 | 65.88 | 67.86 | 66.27 | 65.85 | 1996:Q3 |
| TBILL-SPF | 81.40*** | 82.35*** | 77.38 | 67.47 | 60.98 | 1996:Q3 |
| RGDP-Greenbook | 75.81** | 77.24*** | 75.41 | 81.82*** | 80.00*** | 1980:Q4 |
| UNRATE-Greenbook | 69.60 | 72.58 | 69.92 | 69.67 | 67.77 | 1980:Q4 |
| PGDP-Greenbook | 70.97 | 79.67*** | 77.87** | 81.82*** | 80.00*** | 1980:Q4 |
| CPI-Greenbook | 66.67 | 69.12 | 68.66 | 69.70 | 70.77 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.09 | 64.71 | 69.05 | 67.47 | 71.95 | 1996:Q3 |
| TBILL-SPF | 79.07* | 88.24*** | 84.52** | 80.72 | 79.27 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon SV-AR(1) model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.11: Density forecast accuracy as measured by CRPS, SV-AR(1) specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (SV-AR1 rel.) | 3.11%*** | 7.29%*** | 7.72%*** | 9.02%*** | 7.37%*** | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (SV-AR1 rel.) | 1.69%** | 3.04%** | 4.25%*** | 4.63%** | 3.79%* | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (SV-AR1 rel.) | 0.43% | 0.85% | 0.85% | 1.84% | 2.86%* | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (SV-AR1 rel.) | 2.90%** | 2.48% | 1.87% | 2.14% | 2.70%** | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (SV-AR1 rel.) | 11.91%*** | 14.06%*** | 13.31%*** | 10.74%*** | 8.19%** | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (SV-AR1 rel.) | 3.04%*** | 5.07%*** | 5.63%*** | 6.59%*** | 4.77%*** | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (SV-AR1 rel.) | 1.37%* | 2.30%** | 3.20%** | 3.07%** | 2.11%** | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (SV-AR1 rel.) | 1.48%* | 1.54%** | 2.24%* | 2.54%* | 2.75%* | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (SV-AR1 rel.) | 3.64%* | 2.95%* | 4.05%** | 3.46%* | 3.57% | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using SV-AR(1) rather than FE-CONST; positive numbers indicate improvement of SV-AR(1) over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.12: Forecast error coverage rates, SV-Factor specification: one-standard-deviation bands

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: SV-Factor** | | | | | | |
| RGDP-SPF | 68.38 | 65.19 | 67.16 | 66.92 | 71.97 | 1983:Q4 |
| UNRATE-SPF | 67.88 | 65.44 | 64.44 | 55.97** | 54.14** | 1983:Q4 |
| PGDP-SPF | 71.32 | 66.67 | 67.91 | 69.17 | 64.39 | 1983:Q4 |
| CPI-SPF | 66.28 | 62.35 | 64.29 | 67.47 | 70.73 | 1996:Q3 |
| TBILL-SPF | 77.91* | 78.82** | 70.24 | 63.86 | 53.66 | 1996:Q3 |
| RGDP-Greenbook | 74.19 | 72.36 | 75.41* | 79.34*** | 86.67*** | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 73.39 | 73.98 | 70.49 | 65.29 | 1980:Q4 |
| PGDP-Greenbook | 66.94 | 73.17 | 76.23* | 76.03* | 71.67 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 64.18 | 65.15 | 69.23 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.09 | 64.71 | 69.05 | 67.47 | 71.95 | 1996:Q3 |
| TBILL-SPF | 79.07* | 88.24*** | 84.52** | 80.72 | 79.27 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon SV-Factor model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.13: Density forecast accuracy as measured by CRPS, SV-Factor specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (SV-1FCTR rel.) | 2.98%** | 7.67%*** | 8.20%*** | 9.62%*** | 8.01%*** | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (SV-1FCTR rel.) | 2.95%** | 3.71%* | 4.61%** | 4.38%* | 2.80% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (SV-1FCTR rel.) | 0.65% | 0.73% | 0.92% | 1.75% | 1.96% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (SV-1FCTR rel.) | 3.60%* | 2.12% | 0.89% | 1.16% | 2.32% | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (SV-1FCTR rel.) | 11.34%*** | 13.29%*** | 11.93%*** | 8.92%** | 6.04% | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (SV-1FCTR rel.) | 2.91%** | 4.43%** | 5.56%*** | 6.81%*** | 4.37%* | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (SV-1FCTR rel.) | 1.46% | 2.78% | 4.03%* | 3.50%* | 1.89% | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (SV-1FCTR rel.) | 2.61%** | 2.00%** | 3.13%** | 3.95%** | 3.89%* | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (SV-1FCTR rel.) | 3.70% | 2.09% | 3.25% | 2.37% | 2.46% | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using SV-Factor rather than FE-CONST; positive numbers indicate improvement of SV-Factor over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.14: Forecast error coverage rates, JOINT-SV specification: one-standard-deviation bands

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: JOINT-SV** | | | | | | |
| RGDP-SPF | 77.94*** | 73.33 | 75.37 | 75.94 | 76.52* | 1983:Q4 |
| UNRATE-SPF | 68.61 | 69.12 | 65.19 | 59.70 | 56.39* | 1983:Q4 |
| PGDP-SPF | 78.68*** | 80.74*** | 82.84*** | 78.20** | 79.55*** | 1983:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available). The upper panel provides results based on our proposed multi-horizon JOINT-SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.15: Density forecast accuracy as measured by CRPS, JOINT-SV specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (JOINT-SV rel.) | 2.59%** | 6.17%*** | 6.54%*** | 7.74%*** | 5.67%** | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (JOINT-SV rel.) | 1.42% | 2.45%* | 3.43%* | 3.49% | 2.28% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (JOINT-SV rel.) | 0.34% | −0.29% | 0.06% | 0.96% | 1.70% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using JOINT-SV rather than FE-CONST; positive numbers indicate improvement of JOINT-SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.16: Forecast error coverage rates, JOINT-SV-Factor specification: one-standard-deviation bands

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: JOINT-SV-Factor** | | | | | | |
| RGDP-SPF | 71.32 | 69.63 | 70.15 | 67.67 | 71.21 | 1983:Q4 |
| UNRATE-SPF | 63.50 | 61.03 | 57.78** | 52.24** | 47.37*** | 1983:Q4 |
| PGDP-SPF | 75.74* | 74.81* | 76.12* | 75.19 | 76.52* | 1983:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available). The upper panel provides results based on our proposed multi-horizon JOINT-SV-Factor model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.17: Density forecast accuracy as measured by CRPS, JOINT-SV-Factor specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (JOINT-SV-Factor rel.) | 2.78%** | 6.51%*** | 6.98%*** | 8.43%*** | 6.72%*** | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (JOINT-SV-Factor rel.) | 1.97% | 2.42% | 3.45% | 3.32% | 1.84% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (JOINT-SV-Factor rel.) | 1.65% | 1.01% | 1.08% | 2.10% | 2.72% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using JOINT-SV-Factor rather than FE-CONST; positive numbers indicate improvement of JOINT-SV-Factor over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.18: Forecast error coverage rates, ETA-CONST specification: one-standard-deviation bands

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **Panel A: ETA-CONST** | | | | | | |
| RGDP-SPF | 79.41*** | 77.78** | 77.61* | 78.20* | 77.27* | 1983:Q4 |
| UNRATE-SPF | 74.45 | 77.94** | 74.81 | 74.63 | 69.92 | 1983:Q4 |
| PGDP-SPF | 78.68*** | 77.78*** | 82.09*** | 77.44** | 77.27** | 1983:Q4 |
| CPI-SPF | 72.09 | 63.53 | 61.90 | 62.65 | 62.20 | 1996:Q3 |
| TBILL-SPF | 81.40** | 85.88*** | 82.14* | 79.52 | 70.73 | 1996:Q3 |
| RGDP-Greenbook | 78.23** | 81.30*** | 77.05* | 77.69* | 80.83** | 1980:Q4 |
| UNRATE-Greenbook | 70.40 | 76.61* | 77.24 | 75.41 | 76.86 | 1980:Q4 |
| PGDP-Greenbook | 78.23** | 81.30*** | 81.15*** | 79.34** | 78.33* | 1980:Q4 |
| CPI-Greenbook | 66.67 | 66.18 | 61.19 | 57.58 | 60.00 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 72.99 | 82.35*** | 85.19*** | 87.31*** | 86.47*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.09 | 64.71 | 69.05 | 67.47 | 71.95 | 1996:Q3 |
| TBILL-SPF | 79.07* | 88.24*** | 84.52** | 80.72 | 79.27 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon ETA-CONST model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.19: Density forecast accuracy as measured by CRPS, ETA-CONST specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (ETA-CONST rel.) | $-0.72\%^{***}$ | $1.67\%^{***}$ | $2.43\%^{***}$ | $3.78\%^{***}$ | $2.57\%^{***}$ | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (ETA-CONST rel.) | $-0.22\%$ | $1.26\%^{***}$ | $2.77\%^{***}$ | $3.63\%^{***}$ | $3.37\%^{**}$ | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (ETA-CONST rel.) | $-0.68\%^{***}$ | $0.29\%$ | $0.26\%$ | $1.07\%$ | $2.08\%^{*}$ | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (ETA-CONST rel.) | $-0.09\%$ | $1.04\%$ | $0.88\%$ | $0.90\%$ | $0.99\%$ | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (ETA-CONST rel.) | $-0.75\%^{**}$ | $3.63\%^{***}$ | $4.53\%^{***}$ | $3.90\%$ | $2.80\%$ | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (ETA-CONST rel.) | $-0.56\%^{***}$ | $0.71\%^{**}$ | $0.23\%$ | $1.13\%^{**}$ | $-0.62\%$ | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (ETA-CONST rel.) | $-0.03\%$ | $-0.33\%$ | $0.85\%^{*}$ | $1.45\%^{**}$ | $1.17\%^{*}$ | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (ETA-CONST rel.) | $-0.39\%^{*}$ | $-1.38\%^{***}$ | $-0.70\%$ | $-0.50\%$ | $-0.02\%$ | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (ETA-CONST rel.) | $-0.03\%$ | $0.21\%$ | $0.57\%$ | $0.41\%$ | $0.78\%$ | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using ETA-CONST rather than FE-CONST; positive numbers indicate improvement of ETA-CONST over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.20: Density forecast accuracy as measured by CRPS, ETA-SV vs ETA-CONST specifications

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (SV rel.) | 3.70%*** | 5.93%*** | 5.67%*** | 5.70%*** | 5.14%** | 1983:Q4 |
| (ETA-CONST) | 0.83 | 1.00 | 1.07 | 1.12 | 1.14 | |
| **UNRATE-SPF** | | | | | | |
| (SV rel.) | 1.99%** | 1.58% | 0.81% | −0.20% | −1.15% | 1983:Q4 |
| (ETA-CONST) | 0.08 | 0.17 | 0.25 | 0.33 | 0.42 | |
| **PGDP-SPF** | | | | | | |
| (SV rel.) | 1.70% | 1.13% | 1.57% | 1.54% | 0.93% | 1983:Q4 |
| (ETA-CONST) | 0.50 | 0.56 | 0.59 | 0.62 | 0.67 | |
| **CPI-SPF** | | | | | | |
| (SV rel.) | 2.07% | 1.32% | 0.62% | 0.73% | 1.37% | 1996:Q3 |
| (ETA-CONST) | 0.67 | 1.04 | 1.09 | 1.09 | 1.09 | |
| **TBILL-SPF** | | | | | | |
| (SV rel.) | 12.02%*** | 10.75%*** | 8.87%*** | 6.19%** | 4.18%* | 1996:Q3 |
| (ETA-CONST) | 0.07 | 0.22 | 0.38 | 0.55 | 0.74 | |
| **RGDP-GB** | | | | | | |
| (SV rel.) | 3.13%** | 4.23%*** | 5.40%*** | 5.47%*** | 5.10%** | 1980:Q4 |
| (ETA-CONST) | 0.90 | 1.21 | 1.33 | 1.38 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (SV rel.) | 1.54% | 2.92%** | 2.96%* | 1.72% | 0.38% | 1980:Q4 |
| (ETA-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.49 | |
| **PGDP-GB** | | | | | | |
| (SV rel.) | 1.89% | 2.54%** | 2.68% | 3.37%* | 3.32% | 1980:Q4 |
| (ETA-CONST) | 0.55 | 0.61 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (SV rel.) | 3.30% | 2.87% | 4.02% | 3.19% | 2.93% | 1994:Q4 |
| (ETA-CONST) | 0.55 | 1.10 | 1.21 | 1.23 | 1.21 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using ETA-SV rather than ETA-CONST; positive numbers indicate improvement of ETA-SV over the ETA-CONST case. The bottom row reports the CRPS for the ETA-CONST case. Both ETA-SV and ETA-CONST have been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.21: Forecast error coverage rates: one-standard-deviation bands, FE-VAR(2)-SV specification

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: FE-VAR(2)-SV** | | | | | | |
| RGDP-SPF | 69.85 | 77.04** | 76.87* | 77.44** | 75.76 | 1983:Q4 |
| UNRATE-SPF | 71.53 | 72.06 | 67.41 | 64.18 | 63.91 | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.78** | 81.34*** | 78.20** | 78.79** | 1983:Q4 |
| CPI-SPF | 73.26 | 76.47* | 76.19* | 75.90 | 75.61* | 1996:Q3 |
| TBILL-SPF | 77.91** | 82.35*** | 73.81 | 69.88 | 62.20 | 1996:Q3 |
| RGDP-Greenbook | 75.81** | 77.24*** | 77.05** | 78.51** | 82.50*** | 1980:Q4 |
| UNRATE-Greenbook | 71.20 | 71.77 | 70.73 | 61.48 | 56.20* | 1980:Q4 |
| PGDP-Greenbook | 71.77 | 71.54 | 77.05** | 76.86* | 77.50* | 1980:Q4 |
| CPI-Greenbook | 69.57 | 67.65 | 67.16 | 69.70 | 70.77 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 73.19 | 82.48*** | 85.29*** | 87.41*** | 86.57*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.41 | 65.12 | 69.41 | 67.86 | 72.29 | 1996:Q3 |
| TBILL-SPF | 79.31** | 88.37*** | 84.71** | 80.95 | 79.52 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon FE-VAR(2)-SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.22: Density forecast accuracy as measured by CRPS, FE-VAR(2)-SV specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (FE-VAR rel.) | 1.70% | 3.68%** | 4.32%** | 4.39%* | 1.28% | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (FE-VAR rel.) | 4.52% | 6.94%* | 6.89% | 5.48% | 2.62% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (FE-VAR rel.) | −2.70% | −3.88% | −2.72% | −2.31% | −4.55% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (FE-VAR rel.) | −4.18% | −1.19% | −0.95% | −1.88% | −3.80% | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (FE-VAR rel.) | 18.86%*** | 21.23%*** | 21.96%*** | 20.35%** | 18.52%** | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (FE-VAR rel.) | 2.49% | 4.74%** | 4.87%** | 4.42% | 2.51% | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (FE-VAR rel.) | 4.78% | 3.66% | 2.89% | 1.13% | −0.29% | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (FE-VAR rel.) | −1.01% | −2.80% | −7.42%** | −13.00%** | −18.47%** | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (FE-VAR rel.) | 0.35% | −0.13% | −0.20% | −2.03% | −3.95% | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using FE-VAR(2)-SV rather than FE-CONST; positive numbers indicate improvement of FE-VAR(2)-SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.23: Forecast error coverage rates: one-standard-deviation bands, FE-VAR(5)-SV specification

| Variable | Forecast horizon 0 | 1 | 2 | 3 | 4 | eval. begin |
|---|---|---|---|---|---|---|
| **Panel A: FE-VAR(5)-SV** | | | | | | |
| RGDP-SPF | 69.85 | 75.56* | 76.12* | 78.20** | 75.76 | 1983:Q4 |
| UNRATE-SPF | 71.53 | 70.59 | 67.41 | 64.93 | 61.65 | 1983:Q4 |
| PGDP-SPF | 74.26 | 76.30** | 79.85*** | 75.19 | 78.03** | 1983:Q4 |
| CPI-SPF | 72.09 | 75.29 | 76.19* | 75.90 | 76.83** | 1996:Q3 |
| TBILL-SPF | 79.07** | 82.35*** | 73.81 | 68.67 | 59.76 | 1996:Q3 |
| RGDP-Greenbook | 75.81** | 78.05*** | 77.05** | 78.51** | 83.33*** | 1980:Q4 |
| UNRATE-Greenbook | 72.80 | 70.97 | 70.73 | 63.11 | 58.68 | 1980:Q4 |
| PGDP-Greenbook | 71.77 | 69.92 | 77.05** | 76.86* | 77.50* | 1980:Q4 |
| CPI-Greenbook | 69.57 | 67.65 | 67.16 | 69.70 | 73.85 | 1994:Q4 |
| **Panel B: FE-CONST** | | | | | | |
| RGDP-SPF | 77.94*** | 78.52** | 77.61* | 78.95* | 79.55** | 1983:Q4 |
| UNRATE-SPF | 73.19 | 82.48*** | 85.29*** | 87.41*** | 86.57*** | 1983:Q4 |
| PGDP-SPF | 75.00* | 77.04** | 77.61** | 78.20** | 79.55*** | 1983:Q4 |
| CPI-SPF | 72.41 | 65.12 | 69.41 | 67.86 | 72.29 | 1996:Q3 |
| TBILL-SPF | 79.31** | 88.37*** | 84.71** | 80.95 | 79.52 | 1996:Q3 |
| RGDP-Greenbook | 75.81* | 79.67*** | 74.59 | 76.03 | 75.83 | 1980:Q4 |
| UNRATE-Greenbook | 68.80 | 72.58 | 74.80 | 76.23 | 79.34* | 1980:Q4 |
| PGDP-Greenbook | 75.81* | 78.86** | 75.41 | 76.86* | 72.50 | 1980:Q4 |
| CPI-Greenbook | 65.22 | 66.18 | 62.69 | 57.58 | 58.46 | 1994:Q4 |

Note: The table reports the empirical out-of-sample coverage rates of one-standard-deviation bands. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). The upper panel provides results based on our proposed multi-horizon FE-VAR(5)-SV model. The lower panel provides results based on the FE-CONST model estimated over rolling windows with 60 quarterly observations. Statistically significant departures from a nominal coverage of 68% (as predicted under a normal distribution) are indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

Table A.24: Density forecast accuracy as measured by CRPS, FE-VAR(5)-SV specification

| | Forecast horizon | | | | | |
|---|---|---|---|---|---|---|
| Variable | 0 | 1 | 2 | 3 | 4 | eval. begin |
| **RGDP-SPF** | | | | | | |
| (FE-VAR rel.) | 1.64% | 3.09% | 3.92%$^{**}$ | 4.13%$^{*}$ | 1.33% | 1983:Q4 |
| (FE-CONST) | 0.82 | 1.02 | 1.10 | 1.16 | 1.17 | |
| **UNRATE-SPF** | | | | | | |
| (FE-VAR rel.) | 3.44% | 6.46% | 5.16% | 2.16% | −2.00% | 1983:Q4 |
| (FE-CONST) | 0.08 | 0.17 | 0.25 | 0.34 | 0.43 | |
| **PGDP-SPF** | | | | | | |
| (FE-VAR rel.) | −0.90% | −1.13% | 0.74% | 1.39% | −1.51% | 1983:Q4 |
| (FE-CONST) | 0.50 | 0.56 | 0.60 | 0.63 | 0.68 | |
| **CPI-SPF** | | | | | | |
| (FE-VAR rel.) | −5.03% | −1.59% | −1.95% | −2.74% | −5.08% | 1996:Q3 |
| (FE-CONST) | 0.66 | 1.05 | 1.09 | 1.10 | 1.10 | |
| **TBILL-SPF** | | | | | | |
| (FE-VAR rel.) | 21.63%$^{***}$ | 23.63%$^{***}$ | 24.02%$^{***}$ | 20.89%$^{***}$ | 17.69%$^{**}$ | 1996:Q3 |
| (FE-CONST) | 0.07 | 0.23 | 0.40 | 0.58 | 0.76 | |
| **RGDP-GB** | | | | | | |
| (FE-VAR rel.) | 1.37% | 3.82%$^{*}$ | 4.41%$^{*}$ | 3.85% | 1.72% | 1980:Q4 |
| (FE-CONST) | 0.90 | 1.22 | 1.33 | 1.39 | 1.37 | |
| **UNRATE-GB** | | | | | | |
| (FE-VAR rel.) | 4.43% | 2.65% | 2.16% | 0.30% | −0.53% | 1980:Q4 |
| (FE-CONST) | 0.10 | 0.20 | 0.30 | 0.40 | 0.50 | |
| **PGDP-GB** | | | | | | |
| (FE-VAR rel.) | −0.47% | −0.54% | −5.51% | −10.31%$^{*}$ | −15.85%$^{**}$ | 1980:Q4 |
| (FE-CONST) | 0.54 | 0.60 | 0.62 | 0.65 | 0.69 | |
| **CPI-GB** | | | | | | |
| (FE-VAR rel.) | −0.52% | −0.17% | 0.73% | −2.83% | −5.15% | 1994:Q4 |
| (FE-CONST) | 0.55 | 1.10 | 1.22 | 1.24 | 1.22 | |

Note: The table reports CRPS results for out-of-sample density forecasts. The sample uses predictions made from the date given in the right-most column through 2017:Q4 (and realized forecast errors as far as available), in case of the SPF, and through 2011:Q4, in case of the Greenbook (evaluated against realized data as far as available in both cases). For each variable, the top row reports the relative CRPS calculated as the percentage decrease of the CRPS when using FE-VAR(5)-SV rather than FE-CONST; positive numbers indicate improvement of FE-VAR(5)-SV over the FE-CONST case. The bottom row reports the CRPS for the FE-CONST case, which has been estimated over rolling windows with 60 quarterly observations. Statistical significance of the differences in average CRPS — assessed with a Diebold and Mariano (1995) test — is indicated by *, **, or ***, corresponding to 10, 5, and 1 percent significance, respectively.

# References

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2016), "Common Drifting Volatility in Large Bayesian VARs," *Journal of Business and Economic Statistics* 34, 375-390.

Clark, Todd E. and Francesco Ravazzolo (2015), "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility," *Journal of Applied Econometrics* 30, 551-575.

Cogley, Timothy, and Thomas J. Sargent (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post-World War II U.S.," *Review of Economic Dynamics* 8, 262-302.

Diebold, Francis X., and Roberto S. Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13, 253-263.

Faust, Jon, and Jonathan H. Wright (2008), "Efficient Forecast Tests for Conditional Policy Forecasts," *Journal of Econometrics* 146, 293-303.

——— (2009), "Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset," *Journal of Business and Economic Statistics* 27, 468-479.

Jacobs, Jan P.A.M., Samad Sarferaz, Jan-Egbert Sturm, and Simon van Norden (2015), "Modeling Multivariate Data Revisions," manuscript.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng (2015), "Measuring Uncertainty," *American Economic Review* 105, 1177-1216.

Kishor, N. Kundan, and Evan F. Koenig (2012), "VAR Estimation and Forecasting When Data Are Subject to Revision," *Journal of Business and Economic Statistics* 30, 181-190.

Koop, Gary (2003), *Bayesian Econometrics*, Wiley-Interscience.

Stark, Tom (2013), "SPF Panelists' Forecasting Methods: A Note on the Aggregate Results of a November 2009 Special Survey," manuscript, Federal Reserve Bank of Philadelphia.