Forecasting US GNP Growth: The Role of Uncertainty[1]

Mawuli Segnon[†]
Center for Quantitative Economics, University of Münster,
and
Mark E AG,
Germany

Rangan Gupta[‡]
Department of Economics, University of Pretoria,
South Africa

Stelios Bekiros[§]
Department of Economics, European University Institute,
IPAG Business School, Paris, France,
Florence, Italy

and
Mark E. Wohar[¶]
Department of Economics, University of NE-Omaha, USA;
School of Business and Economics, Loughborough University, UK
October 11, 2017

Abstract: A large number of models have been developed in the literature, in order to analyze and forecast the changes in output dynamics. The objective of this paper is to compare the predictive ability of uni- and bivariate models, in terms of forecasting US GNP growth at different forecasting horizons, with the bivariate models containing information on a measure of economic uncertainty. Based on point and density forecast accuracy measures, as well as on equal predictive ability (EPA) and superior predictive ability (SPA) tests (cf. Diebold and Mariano, 1995; Hansen, 2005), we evaluate the

---

[†]Corresponding author: Mawuli Segnon, Center for Quantitative Economics, Am Stadtgraben 9, 48143 Münster, Germany. E-mail: segnon@uni-muenster.de.

[‡]Rangan Gupta, Department of Economics, University of Pretoria, South Africa.
E-mail: rangan.gupta@up.ac.za.

[§]Stelios Bekiros, Department of Economics, European University Institute, Florence, Italy.
E-mail: stelios.bekiros@eui.eu.

[¶]Mark E. Wohar, Department of Economics, University of NE-Omaha, USA,
School of Business and Economics, Loughborough University, UK. E-mail: mwohar@unomaha.edu.

relative forecasting performance of different model specifications over the quarterly period of 1919:2 until 2014:4. We find that the economic policy uncertainty (EPU) index should improve the accuracy of US GNP growth forecasts in bivariate models. We also find that the EPU exhibits similar forecasting ability, as the term spread and outperforms other uncertainty measures such as the volatility index and geopolitical risk in predicting US recessions. While the Markov-switching time varying parameter vector autoregressive (MS-TVP-VAR) model yields the lowest values for the root mean squared error (RMSE) in most cases, we observe relatively low values for the log predictive density score (LPDS), when using the Bayesian VAR (BVAR) model with stochastic volatility. More importantly, our results highlight the importance of uncertainty in forecasting US GNP growth rates.

1 Introduction

Theoretical papers by Bloom (2009), Mumtaz and Zanetti (2013) and Carriero et al. (2015), following the early works of Bernanke (1983), Dixit and Pindyck (1994), confirm that, besides productivity and/or policy shocks, various forms of policy generated uncertainty lead to business cycle fluctuations. While the (negative) influence of uncertainty on economic activity is well-established theoretically, in the wake of the "Great Recession", the focus has also been on quantifying the impact of uncertainty. Understandably, this requires a measure of uncertainty - an otherwise latent variable. In this regard, there are two approaches to measuring uncertainty: (1) The news-based approach of Brogaard and Detzel (2015) and Baker et al. (2016), whereby the authors perform month-by-month searches of newspapers for terms related to economic and policy uncertainty to construct their measure of Economic Policy Uncertainty (EPU); (ii) Alternatively, Mumtaz and Zanetti (2013), Mumtaz and Surico (2013), Alessandri and Mumtaz (2014), Mumtaz and Theodoridis (2015, 2016), Bali et al. (2015), Carriero et al. (2015), Chuliá et al. (2015), Jurado et al. (2015), Ludvigson et al. (2015), Rossi and Sekhposyan (2015), Rossi et al. (2016), Shin and Zhong (2016), Creal and Wu (2017) recover measures of uncertainty from the estimation of various types of small- and large-scale structural models related to macroeconomics and finance. Irrespective of which approach (news- or model-based) is pursued, these studies, along with others that have used such indices (for example, Bachmann and Bayer (2011), Knotek II and Khan (2011), Bachmann et al. (2013), Colombo (2013), Jones and Olson (2013), Benati (2013), Caggiano et al. (2014), Kang et al. (2014), Karnizova and Li (2014), Castelnuovo et al. (2015), Cheng et al. (2016), and Balcilar et al. (2016)) confirm the significant role of uncertainty in affecting economic activity.

However, apart from Karnizova and Li (2014), and Balcilar et al. (2016), all the above-mentioned studies trying to link uncertainty with economic activity (for example, measures of output and/or unemployment, investment) have entailed in-sample analysis. Karnizova and Li (2014) depict the role that the news-based EPU of Baker et al. (2016) can play in forecasting US recessions based on probit models. By contrast, Balcilar et al. (2016) emphasize that forecasting gains for US recessions can be obtained using mixed-frequency Markov-switching models. Against this backdrop of limited evidence on out-of-sample forecasting of a measure of economic activity, and given the widely held view that importance of variables and models require out-of-sample validation (Campbell, 2008), the objective of our paper is to use a wide-array of univariate and multivariate linear and nonlinear models in analysing the role played by the news-based measure of EPU of Baker et al. (2016) in forecasting US Gross National Product (GNP) growth rate. In our study, we analyze the forecasting performances of the various models considered over the historical quarterly period of 1919:2 to 2014:4, using an in-sample of 1900:1 to 1919:1. The decision to use the news-based EPU, rather than model-based uncertainty, simply emanates from the availability of a measure of uncertainty for forecasting GNP growth over the longest possible sample period, covering various phases of

the US economic history.

Forecasts of output growth represent an important indicator for both policymakers and financial investors. Such figures reveal information about the current state of the economy and play a key role in formulating appropriate monetary and fiscal policies. As indicators of future growth potential of the economy, they help financial investors in their investment decision making process. Hence, the need for accurately forecasting the growth rate of the economy cannot be overstated. Given the importance of forecasting economic growth, different powerful uni- and multivariate econometric models have been developed in the literature to provide accurate GDP growth forecasts, especially of the vector autoregressive (VAR) variety (cf. Rossi and Sekhposyan, 2010, 2014; Eickmeier et al., 2011; Schumacher, 2011; Chauvet and Potter, 2013; Giannone et al., 2015; Schorfheide and Song, 2015, for an overview of different models for forecasting output growth). This paper considers the baseline ARMA model, the Bayesian VAR (BVAR), the threshold VAR (TVAR), the smooth transition VAR (ST-VAR), two-types of time varying parameter VARs (TVP-VARs), the Markov switching VAR (MSVAR), the unobserved component stochastic volatility (UCSV), the Bayesian VAR with CSV and a Mixed-frequency VAR (MF-VAR) to produce both point and density forecasts of U.S. GNP growth. As discussed in Rossi and Sekhposyan (2010), Bekiros and Paccagnini (2013) and D'Agostino et al. (2013), it is important to model nonlinearities when forecasting US output, due to issues of structural instability, and also when relating movements of output with uncertainty (Caggiano et al., 2014). Hence, we look at both linear and nonlinear models. In addition, Herbst and Schorfheide (2012), Barnett et al. (2014), Rossi and Sekhposyan (2014) argue that it is becoming more and more important to assess the uncertainty associated with the forecasts of models. This is specifically why central bankers wish to evaluate how well models perform in forecasting a range (uncertainty) of future values of relevant macroeconomic variables, rather than just the point forecasts of these variables. In other words, the forecaster needs to look not only at point forecasts, but also analyze density forecasts (Bekiros and Paccagnini, 2015a). Note that in this paper, we basically take an atheoretical approach, although there are of course theoretical models of forecasting output based on large-scale Keynesian-type models and microfounded Dynamic Stochastic General Equilibrium models (cf. Bekiros and Paccagnini, 2013, 2014, 2015b; Del Negro and Schorfheide, 2012; Del Negro et al., 2016, for detailed reviews in this regard). It must be emphasized that our objective in this paper is not necessarily to contribute to the model sets used in forecasting output growth. Rather, the objective is primarily that of forecasting US GNP growth, based on existing models, but for the first time, incorporating the role of EPU.

The remainder of the paper is organized as follows. Section 2 presents the data used in our analysis. In Section 3, we describe the different forecasting models used in this study. Section 4 provides the forecasting evaluation methodologies. The empirical results are presented in Section 5 and Section 6 concludes.

2 Data analysis

This study uses time series of different frequency that stem from different sources. The quarterly data on real US GNP covering the time period 1900:1 to 2014:4 is obtained from Omay et al. (2016), who combine two different sources to create a long span of these observations. First, they collect the observations covering the time period of 1900:1 to 1946:4 from Gordon (1986).[6] Second, the pre-1947 data is completed using observations available at the FRED database of the Federal Reserve Bank of St. Louis.

The monthly EPU measure used in this study corresponds to the historical measure of uncertainty for the US economy recently developed by Baker et al. (2016). The uncertainty index is created from two overlapping sets of ten newspapers (the Wall Street Journal, New York Times, Washington Post, Chicago Tribune, LA Times, Boston Globe, USA Today, Miami Herald, Dallas Morning Tribune, and the San Francisco Chronicle). Starting in January, 1900 until December, 1985, for each month, they collect articles published in the first six above-mentioned newspapers whose information contents are nevertheless fit into the following three categories: uncertainty, economy and policy. Each article published in a paper within a specific month is selected, under the condition that it, at the same time, contains the term *uncertainty* or *uncertain*, the terms *economic*, *economy*, *business*, *commerce*, *industry* and *industrial* and one or more of terms such as: *congress*, *legislation*, *white house*, *regulation*, *federal reserve*, *deficit*, *tariff*, or *war*. From January, 1985 to December 2014, the same search described above is performed each month, based on the articles published in all ten newspapers. A normalization procedure is performed to overcome the fluctuation observed in the volumes of news articles for a predetermined newspaper over time. Here, we do not describe the normalization procedures in detail because it is beyond the scope of our study. We refer the reader to Baker et al. (2016) for more details on the normalization procedures. The data are available for download from https://www.policyuncertainty.com/us_monthly.html. Note that we compute quarterly values of the EPU index by taking three-months averages to come up with a quarterly value for this index. The monthly value of the index is used for the mixed frequency model of forecasting (discussed below in detail).

We compute the percentage US GNP growth, g_t, as follows

$$g_t = 100 * \log\left(\frac{\text{GNP}_t}{\text{GNP}_{t-1}}\right), \tag{1}$$

where GNP_t denotes the real US GNP at the time t.

The descriptive statistics of the percentage US GNP growth and the logarithmized EPU are reported in Table 1. We applied the augmented Dickey-Fuller (ADF) tests and the results in Table 1 reject the null hypothesis of unit root for both the percentage US GNP growth and the logarithmized EPU. These results indicate that the natural logarithm of EPU is stationary, and hence, we work with the series in its log-level form. We observe

---

[6]The original source of the data by Gordon (1986) is the National Bureau of Economic Research (NBER).

a higher variability in the real US GNP growth than in the logarithmized EPU. *Fig.* 1 depicts the quarterly US real GNP (in billions), the corresponding quarterly growth (as defined in *eq.* (1) and the logarithmized EPU. The vertical bars in *Fig.* 1 depict the National Bureau of Economic Research (NBER) dated-recessions.

3 Forecasting models

Different model specifications have been developed in the literature for forecasting output growth. In this section, we briefly describe the uni- and bivariate econometric models that we use in our forecasting exercises.

3.1 Univariate models

3.1.1 Autoregressive moving average model

The autoregressive moving average (ARMA) model with constant shock variance is the most popular and commonly used univariate model in forecasting output growth. The ARMA(p,q) can be formalized as

$$g_t = c + \sum_{i=1}^{p} a_i g_{t-i} + \sum_{j=1}^{q} b_j \epsilon_{t-j} + \epsilon_t, \tag{2}$$

where g_t is US GNP growth and ϵ_t denotes innovation in the model and is assumed to follow a normal distribution ($\epsilon_t \sim N(0, \sigma)$).

3.1.2 Unobserved component model with stochastic volatility

Another variant of the univariate model is the unobserved component model with stochastic volatility (UCSV). Studies by Stock and Watson (2007); Barnett et al. (2014) show that the UCSV is appropriate for modeling and forecasting output growth and inflation rates. As in Stock and Watson (2007), the UCSV can be expressed as

$$
\begin{aligned}
g_t &= \mu_t + u_t, && \text{with } u_t = \sigma_{u,t}\xi_{u,t} \\
\mu_t &= \mu_{t-1} + v_t, && \text{with } v_t = \sigma_{v,t}\xi_{v,t},
\end{aligned}
\tag{3}
$$

where g_t is the US GNP growth, μ_t denotes the stochastic trend in the model and the innovation vector $\xi_t = (\xi_{u,t}, \xi_{v,t}) \sim iid\ N(0, I)$. The logaritmized variances $\sigma_{u,t}$ and $\sigma_{v,t}$ in *eq.* (3) evolve as independent random walks:

$$
\begin{aligned}
\ln \sigma_{u,t}^2 &= \ln \sigma_{u,t-1}^2 + \varepsilon_{u,t} \\
\ln \sigma_{v,t}^2 &= \ln \sigma_{v,t-1}^2 + \varepsilon_{v,t},
\end{aligned}
\tag{4}
$$

where $\varepsilon_t = (\varepsilon_{u,t}, \varepsilon_{v,t}) \sim iid\ N(0, \lambda I)$ and ξ_t and ε_t are independently distributed.

3.2 Bivariate models

3.2.1 Bayesian vector autoregressive (BVAR) model

To obtain more accurate output growth forecasts the vector autoregressive (VAR) models have been developed in the literature. VAR models enable the use of all available information, e.g. leading indicators at the time used to produce the forecast. A baseline VAR(p) model with constant variance-covariance of shocks has the following form

$$x_t = \Phi_1 x_{t-1} + \cdots + \Phi_p x_{t-p} + c + \xi_t, \tag{5}$$

where x_t is a $T \times 2$ data matrix that contains the GNP growth and logarithmized EPU.

By defining $z_t = [x'_{t-1}, \ldots, x'_{t-p}, 1]'$ and $\Phi = [\Phi_1, \ldots, \Phi_p, c]'$, the baseline VAR(p) can be expressed as

$$g_t = Z'_t \phi + \xi_t, \tag{6}$$

where $Z_t = I_n \otimes z_t$, $\phi = vec(\Phi)$, and ξ_t is a vector of Gaussian random variables with the covariance matrix Σ. The baseline VAR model is referred to as Bayesian VAR (BVAR), when it is estimated with Bayesian methods (cf. Koop and Korobilis, 2010). Koop and Korobilis (2010) use a variety of priors and estimation methods for BVARs and do not find significant differences in the estimation results. As in Koop and Korobilis (2010) and Del Negro and Schorfheide (2011) we also use Minnesota priors that enable simple posterior and predictive forecasts.

3.2.2 BVAR with Common Stochastic Volatility Model

Recent studies by Clark (2011) and Carriero et al. (2015) show that a combination of the BVAR with common stochastic volatility, brings about an improvement in the forecasting performance of these models. Following Carriero et al. (2015), the BVAR with common stochastic volatility (BVAR-CSV) model can be formulated as

$$
\begin{aligned}
x_t &= Z'_t \phi + \xi_t, \\
\xi_t &= D^{-1} f_t^{0.5} v_t, &\quad v_t &\sim iid\ N(0, I) \\
\ln f_t &= \psi \ln f_{t-1} + u_t, &\quad var(u_t) &= \phi,
\end{aligned} \tag{7}
$$

where x_t is a $T \times 2$ data matrix that consists of the GNP growth and logarithmized EPU, D^{-1} is a lower triangular matrix, and f_t is a scalar process.

3.2.3 Threshold VAR Model

In contrast to the linear VARs, the threshold VAR (TVAR) model allows for capturing a nonlinear dependence structure between macroeconomic variables (cf. Wolters et al.,

1998; Avdjiev and Zeng, 2014) and is defined as follows

$$
\begin{aligned}
x_t &= \Phi_{11}x_{t-1} + \cdots + \Phi_{1p}x_{t-p} + c_1 + \xi_{1t}, & var(\xi_{1t}) = \Omega_1 \text{ if } x_{t-d} \le x_1^* \\
x_t &= \Phi_{21}x_{t-1} + \cdots + \Phi_{2p}x_{t-p} + c_2 + \xi_{2t}, & var(\xi_{2t}) = \Omega_2 \text{ if } x_{t-d} > x_1^*,
\end{aligned}
\tag{8}
$$

where x_t is a $T \times 2$ data matrix that consists of the GNP growth and logarithmized EPU, and x_{t-d} denotes d^{th} lag of GNP growth, and x_1^* is the threshold level of growth that indicates expansions or recessions. The delay parameter means that if the threshold variable x_{t-d} outruns the threshold level x_1^* at time $t - d$, the dynamics actually change at time t.

3.2.4 Smooth transition VAR model

The smooth transition VAR (ST-VAR) model has the following form

$$
x_t = \sum_{i=1}^{p} \Phi_{1i}x_{t-i} + c_1 + \pi\left(\lambda, x_1^*, x_{t-d}\right)\left(\sum_{i=1}^{p} \Phi_{2i}x_{t-i} + c_2\right) + \xi_t,
\tag{9}
$$

where x_t is a $T \times 2$ data matrix that consists of the GNP growth and logarithmized EPU and $\pi(\cdot)$ is a logistic transition function given by

$$
\pi\left(\lambda, x_1^*, x_{t-d}\right) = \left[1 + \exp\left(-\lambda(x_{t-d} - x_1^*)\right)\right]^{-1}
\tag{10}
$$

with $\lambda > 0$ the smoothing parameter. x_1^* is the threshold value around which the dynamics of the model change.

3.2.5 Markov switching VAR (MS-VAR) model

One of the merits of the Markov switching VAR model is that it accounts for the possibility of structural shifts in the data. The MS-VAR framework allows the parameters of the underlying data-generating process of the observed time series to be conditioned on the latent regime variable δ_t. The MS-VAR model is given by

$$
x_t = c_{\delta_t} + \sum_{i=1}^{p} \Phi_{\delta_t}X_{t-i} + \xi_t, \qquad var(\xi_t) = \Omega_{\delta_t},
\tag{11}
$$

where x_t is a $T \times 2$ data matrix consisting of the GNP growth and logarithmized EPU and, Φ_{δ_t} and Ω_{δ_t} are regime-dependent autoregressive coefficients and variance-covariance matrices. δ_t is a discrete process taking its values in $[1, S]$. δ_t is the latent variable that controls the state of the economy, which can be equal to 1,2,..., or S, with S the number of states in the economy. Here, we assume that S is equal to 2. We note that δ_t is a first-order Markov chain that is characterized by the following transition

probabilities p_{ij} between the different states of the economy:

$$p_{ij} = Pr(\delta_t = j | \delta_{t-1} = i), \text{ with } \sum_{j=1}^{S} p_{ij} = 1, \forall i, j \in \{1, \dots, S\}. \tag{12}$$

3.2.6 Time-varying parameter VAR (TVP-VAR) model

By allowing the parameters to change over time, the TVP-VAR represents an appropriate modeling approach that can take into account the economic dynamics that evolve over time (cf. Cogley and Sargent, 2002; Primiceri, 2005 for original contributions to the development of the TVP-VAR model and Nakajima, 2011 for a detailed overview of the methodology and empirical applications). Following Del Negro and Schorfheide (2011) the TVP-VAR model can be formalized as

$$x_t = Z_t' \phi_t + \xi_t, \qquad var(\xi_t) = \Sigma_t \tag{13}$$

where x_t is a $T \times 2$ data matrix that consists of the GNP growth and logarithmized EPU and the parameters, ϕ_t, evolves according to the random walk process

$$\phi_t = \phi_{t-1} + e_t, \qquad e_t \sim iidN(0, Q). \tag{14}$$

The covariance matrix Q is restricted to be diagonal and e_t are uncorrelated with ξ_t. The innovations ξ_t are normally distributed with variance covariance matrix Σ_t.

$$\xi_t \sim N(0, \Sigma_t), \qquad \Sigma_t = D_t^{-1} H_t (D_t^{-1})', \tag{15}$$

where D_t is a lower triangular matrix and H_t is a diagonal matrix whose elements, $h_{i,t}^2$ follow a geometric random walk:

$$\ln h_{i,t} = \ln h_{i,t-1} + \eta_t, \qquad \eta_t \sim iidN(0, \sigma_i^2). \tag{16}$$

3.2.7 Markov-switching time-varying parameter VAR (MS-TVP-VAR) model

We now extend the TVPVAR model to the Markov-switching time-varying parameter VAR (MS-TVP-VAR) model, by allowing the time varying parameters to be dependent on an unobservable variable δ_t that controls the state of the economy (cf. Bekiros and Paccagnini, 2013, 2015b, for original contributions to the development of this framework and applications for forecasting). The MS-TVP-VAR has the following form:

$$x_t = Z_t' \phi_{t,\delta_t} + \xi_t, \qquad var(\xi_t) = \Sigma_t, \tag{17}$$

where x_t is a $T \times 2$ data matrix that consists of the GNP growth and logarithmized EPU and ϕ_{t,δ_t} is a time-varying regime-dependent autoregressive coefficient.

3.2.8 Mixed frequency VAR (MF-VAR) model

Recently developed by Schorfheide and Song (2015), the MFVAR model is a useful tool that allows the modeling and forecasting of macroeconomic variables with different frequencies

$$x_t = \Phi_1 x_{t-1} + \cdots + \Phi_p x_{t-p} + c + \xi_t, \qquad \xi_t \sim iidN(0, \Sigma) \tag{18}$$

where $x_t = [x'_{m,t}, x'_{q,t}]$. The vector $x_{m,t}$ consists of variables that are observed with a monthly frequency, for example, the economic policy uncertainty index, and the vector $x_{q,t}$ contains the unobserved monthly variables that are only released quarterly (for example, US GNP growth).

4 Forecasting evaluation methodologies

We estimate our portfolio of models using a rolling forecasting scheme. We start with observations from 1900:1 to 1919:1 as in-sample and use those from 1919:2 to 2014:4 as out-of-sample. Our rolling forecasting scheme consists of removing one earlier observation and adding a new observation quarter by quarter, so that the estimation sample size remains constant over the out-of-sample period. For each iteration, we produce US GNP growth forecasts up to 8 quarters ahead. All the models are estimated using Bayesian Markov Chain Monte Carlo (MCMC) methods. Using a Gibbs Sampler, draws from the posterior distribution can be generated, and based on these draws, future trajectories of x_t can be simulated to characterize the predictive distribution related to each model specification. We first evaluate the performance of our models via two well-known measures, namely the root mean squared error (RMSE) and the log predictive density score (LPDS).

One of the most popular univariate measures of point forecast accuracy is the root mean squared error that is given by

$$\text{RMSE}_h = \sqrt{\frac{1}{n-h} \sum_{t=1}^{n-h} (\hat{x}_{t+h} - x_{t+h})^2}, \tag{19}$$

where $\hat{x}_{t+h}$ denotes the US real GNP growth forecast at time $t + h$ obtained using univariate or bivariate models and x_{t+h} is the actual quarter US real GNP growth. $(n - h)$ denotes the number of evaluated h-step-ahead forecasts. Small RMSE values indicate good forecasting performance

However, RMSE provides only information on how well a specific model captures the dynamics of GNP growth around the mean. To obtain more information and examine the distribution of our forecasts, we apply the log predictive density score (LPDS) proposed in Adolfson et al. (2007). Formally, we have

$$\text{LPS}_h = -2 \log p_t(x_{t+h}), \tag{20}$$

where $p_t(x_{t+h})$ is the h-step-ahead forecast distribution of the $n-$dimensional data vector x_{t+h}.

Following Adolfson et al. (2007), we also assume that $p_t(x_{t+h})$ is normally distributed and the log predictive density score can be rewritten as

$$\text{LPS}_h = \left(n\log(2\pi) + \log|Q_{t+h|t}| + (x_{t+h} - \bar{x}_{t+h|t})^{'} Q_{t+h|t}^{-1} (x_{t+h} - \bar{x}_{t+h|t}) \right) \tag{21}$$

x_{t+h} is the observed outcome, $\bar{x}_{t+h|t}$ denotes the posterior mean of the forecast distribution and $Q_{t+h|t}$ is the posterior variance of the forecast distribution.

To conclude whether a specific model relatively outperforms others, we make use of the equal predictive ability (EPA) and superior predictive ability tests recently developed by Diebold and Mariano (1995) and Hansen (2005), respectively. In the following sections, we briefly describe both test procedures.

4.1 Equal predictive ability test

Proposed by Diebold and Mariano (1995), the EPA test allows us to compare the forecasting performance of two competitive models (say M_1 and M_2) under a predefined loss function. The null hypothesis states that there is no difference between forecasts from the two models and can be formalized as

$$H_0: \quad E(d_{t,h}) = 0 \quad \forall t, h, \tag{22}$$

where $d_{t,h} = L(\xi_{t,h,M_1}) - L(\xi_{t,h,M_2})$, $\xi_{t,h,M_i} = \hat{x}_{t,h,M_i} - x_{t,h}$ for $i = 1, 2$ and $L(.)$ is the predefined loss function.

The corresponding EPA test statistic is defined as

$$\text{EPA}_h = \frac{\bar{d}_h}{\sqrt{\frac{1}{T} \sum_{k=-N}^{N} \hat{\gamma}_h(k)}}, \tag{23}$$

where $\bar{d}_h = \frac{1}{T} \sum_{T}^{t=1} d_{t,h}$, N is the nearest integer larger than $T^{1/3}$, h is the forecasting horizon and $\hat{\gamma}_h(k)$ can be formalized as

$$\hat{\gamma}_h(k) = \frac{1}{T} \sum_{t=|k|+1}^{T} (d_{t,h} - \bar{d}_h)(d_{t-|k|,h} - \bar{d}_h).$$

Diebold and Mariano (1995) show that under the null hypothesis and in large samples the EPA test statistic approximately follows a normal distribution and the p-values of the EPA test can easily be obtained.

4.2 Superior predictive ability test

In order to compare the relative performance of a particular model with its competitors, based on a pre-specified loss function, Hansen (2005) proposes the superior predictive ability (SPA) test. This test is designed using the framework of the reality check test developed by White (2000). We test the null hypothesis that the benchmark model outperforms any of the other competitive models at the forecasting horizon h. Formally, we have:

$$H_0 : \max_{i=1,\dots,Q} \mathbb{E}\left[d_{t,h}\right] \leq 0, \tag{24}$$

where $d_{t,h} = (d_{i,t,h}, \dots, d_{K,t,h})'$ is a vector of differences between the loss function of the benchmark model and those of the competitive models. $d_{i,t,h}$ are calculated as $d_{i,t,h} = L_{t,h}(M_0) - L_{t,h}(M_i)$, where $L_{t,h}(M_0)$ and $L_{t,h}(M_i)$ are the loss functions at time t and forecast horizon h for a benchmark model M_0 and for its competitor models, $M_{i_{(i=1,\dots,Q)}}$, respectively. Q is the number of competitive models.

The test statistic related to the SPA is given by

$$\text{SPA}_h = \max_{i=1,\dots,K} \frac{\sqrt{T}\bar{d}_{i,h}}{\sqrt{\lim_{T\to\infty} \mathbb{V}ar(\sqrt{T}\bar{d}_{i,h})}}, \tag{25}$$

where $\bar{d}_{i,h} = T^{-1} \sum d_{t,h}$. The p-values of the SPA_h are obtained via s stationary bootstrap procedure.[7]

5 Forecasting results

For each model specification, we compute the root mean squared errors (RMSEs) and the log predictive density scores (LPDS) for the out-of-sample period 1919:2-2014:4 and for different forecasting horizons. Note that the out-of-sample period we consider in this study covers different eras of history that are characterized by economic expansions or recessions. It includes the era of post World War I, the severe recessions of 1929-1933, the post World War II era, the oil price shocks of the 1970s, the deep recession in the early 1980s, the long expansions of the 1960s, 1980s and 1990s, and the recent Global Financial crisis of 2007-2008 and the following Great Recessions. The results are reported in Tables 2 and 3, respectively.

Except for the STVAR and TVPVAR models, the bivariate models in most cases outperform the univariate models. This suggests that the economic policy uncertainty index has information content that helps improve the accuracy of the US GNP growth fore-

---

[7]More information on technical issues and the framework of the SPA test are available in Hansen (2005).

casts. According to the RMSEs, the MSTVPVAR seems to be the best model, as it has the lowest RMSE and cannot be outperformed by other competing models. At the 1-quarter forecast horizon, the RMSE has its lowest value (2.184) relative to longer horizon forecasts. The highest RMSE is at quarter 5 with a value of 2.734.

Fig. 3 depicts the forecasts obtained from the MSTVPVAR model for all forecasting horizons (h=1,...,8) and the actual values of US real GNP growth, and *Fig.* 4 depicts those from the BVAR model with stochastic volatility and the actual values of US real GNP growth. At the 1- and 2- quarter forecast horizons, both models provide relatively good forecasting performance. The difference between the forecasts and the actual data for both models are not so pronounced. At the 3-quarter forecast horizon and beyond, while we observe that the forecasts from the BVAR model with stochastic volatility sometimes run parallel to the actual data, and the deviations from the actual data become larger and larger as the forecast horizon increases, the forecasts from the MSTVPVAR model more accurately approximate the actual data.

The MSTVPVAR and BVAR with stochastic volatility models produce out-of-sample forecasts the behavior of which seems linked to business-cycle fluctuations in *Figs.* 3 and 4. During post World War I recessions, the mid-1970s, the early 1980s, and the 1990-1991 recessions and in particular, the severe Great recessions of 1929-1933 and of 2008-2009 the forecasts identify the local minima during the course of recessions, first remaining very low and then increasing steadily during the expansion phases. This pattern supports the economic plausibility of the use of EPU in forecasting US GNP growth.[8] It seems that the MSTVPVAR and BVAR with stochastic volatility models provide a good data fit.

According to the log predictive density score (LPDS) at the 1-quarter and 8-quarter horizons, the BVAR exhibits the smallest LPDS, followed by the MSVAR at the 2-quarter horizon. The Bayesian VAR model with stochastic volatility (BVARCSV) provides, in most cases the most accurate forecasts and improves compared to other models, at the 3-quarter up to 7-quarter horizon, with the LPDS values being between 4.1 and 4.2. It should be noted that the BVAR and TVAR models have LPDS scores of 4.14 to 4.27 for 3- to 7-quarter forecast horizons, with values close to BVARCSV.

Based on the SPA test results (Table 4) we see that the MSTVPVAR model, followed by the TVAR, BVARCSV and BVAR, cannot be outperformed by other competing models at the 10% confidence level for all forecasting horizons. At the 2-quarter ahead horizon and beyond, the MFVAR and ARMA models also perform well and dominate other models. For the UCSV and the STVAR models, the p-values of the SPA test for all forecasting horizons are smaller than 10%, suggesting the rejection of the null hypothesis at the 10% confidence level. While the UCSV performs worst, the baseline ARMA model cannot be rejected at the 10% confidence level at the 2-quarter horizon and beyond.

We also compare our best models to the remaining ones, and the results are presented

---

[8]Similar patterns can be observed using the term spread in bivariate models.

in Tables 5 and 6. The pairwise comparison between the MSTVPVAR and other bi-variate models (BVAR, TVAR, STVAR, MSVAR, TVPVAR, and MFVAR), based on the root mean squared error, shows that the forecasts from MSTVPVAR model for the 1-quarter ahead horizon up to the 3-quarter ahead horizon are superior to those of other bivariate models at the 10% confidence level. At the 4-quarter horizon and beyond, all the bivariate models perform equally well, cf. Table 5.

When applying the equal accuracy test of Diebold and Mariano (1995) to the BVARCSV and the remaining competitive models, we obtain a clear superiority of the BVARCSV over the STVAR and TVPVAR at the 2-quarter ahead horizon and beyond, and of the MSTVPVAR for all forecasting horizons and the MFVAR at the 1-quarter ahead horizon, cf. Table 6.

To assess how the EPU performs in different model specifications over different eras that are covered by our out-of-sample period, we compute the cumulative difference in squared forecast errors for our best models in point and density forecasts (MSTVPVAR and BVARCSV, respectively) vis-á-vis the remaining competitive models. Figures 5, 6, 7 and 8 show the cumulative difference in squared forecast errors for EPU and term spread at 1- and 2-quarter forecasting horizon for MSTVPVAR and BVARCSV relative to the competitors.[9] The advantage of using this graphical tool is that it is highly in-formative and can help us to compare the mean squared error of our best models to the competitors for any period. All the lines in *Figs.* 5 and 7 have negative slopes or are flat, with relatively infrequent short time periods that have a positive slope. This observation indicates that the MSTVPVAR model cannot be outperformed by the competitors during these time periods. In *Figs* 6 and 8, the tendency is quite clear. All the lines are neg-atively sloped, suggesting a clear superior forecasting performance of the BVARCSV model over the competitors.

5.1 Comparison with other predictors

In this section we compare the out-of-sample forecasting performance of EPU with other leading predictors, namely the term spread, historical data on volatility index (NVIX) and geopolitical risk (GPR). The out-of-sample performance of the term spread in pre-dicting US GDP growth or recessions is well documented in the literature (see Wheelock and Wohar (2009) for a recent review on the predictive power of the term spread). The data on the term spread are available at a quarterly frequency, and those on NVIX and GPR at a monthly frequency covering the time period from January 1900 until Decem-ber 2014. We transform the monthly data into quarterly ones by taking the average over a quarter. *Fig* 2 depicts the quarterly data on the term spread, NVIX and GPR with US observed NBER-dated recessions.

---

[9]These observations also hold for the remaining forecasting horizons

5.1.1 Forecasting US GNP growth using the term spread and other uncertainty measures

In this section, we compare the predictive ability of the EPU with that of the term spread. We produce out-of-sample GNP growth forecasts using the term spread as the predictor in the bivariate models. The out-of-sample forecasts are obtained via rolling forecasting exercises, as explained in Section 4. The root mean squared errors (RMSEs) and the log predictive density scores (LPDS) at different forecasting horizons are reported in Tables 2 through 6. In most cases, we obtain similar results. It seems that the MS-TVPVAR and BVARCSV models cannot be outperformed by other competitive models. We note that for the historical data on NVIX and GPR, we only run our best models in point and density forecasts, i.e. the MS-TVPVAR and BVARCSV models. The out-of-sample results suggest that EPU and the term spread have similar predictive power in the bivariate models and outperform NVIX and GPR.

5.1.2 Predicting US recessions

In this section, we examine the out-of-sample performance of historical EPU, NVIX and GPR in predicting US recessions. We use a static probit model that can easily be estimated via the maximum likelihood method. Following Estrella and Mishkin (1998), we estimate the following equation:

$$Pr(R_{t+h} = 1) = \Phi(\varsigma + \beta x_t),$$

where x_t is a vector of the independent variable such that either EPU, or NVIX, or GPR in the case of a one-factor model. In the case of a multi-factor model, x_t is a matrix of the independent variables: EPU, NVIX and GPR. R_{t+h} denotes the observable US recession indicator at time $t + h$, which is 1 if there is a recession, otherwise 0. Φ is the cumulative normal distribution.

To assess the robustness of the predictive power of EPU, NVIX and GPR, we consider two different time periods: The first covers from 1900:1 to 2014:4 and the second from 1985:1 to 2014:4. We perform rolling out-of-sample forecasting exercises using one-factor and multi-factor models for each time period. For the first period, we use data from 1900:1 to 1919:1 as in-sample and the remaining observations for the period 1919:2 to 2014:4 as out-of-sample, and for the second period we utilize data from 1985:1 to 2004:4 as in-sample and compute out-of-sample for the period 2005:1 to 2014:4. We make use of the pseudo R^2 developed by Estrella (1998) to evaluate the out-of-sample predictive power of each variable, EPU or NVIX or GPR and EPU, NVIX and GPR together. The advantage of pseudo R^2 is that it can easily be computed and enables to ranking different predictors in terms of their explanatory power. The pseudo R^2 values reported in Table 7 indicate that EPU and NVIX exhibit similar explanatory power, but less predictive power than GPR, and EPU, NVIX and GPR together, for the time period 1919:2 to 2014:4. However, the pseudo R^2 values for all variables (one- or multi-factor model) are

very close to zero, indicating little explanatory power in predicting US recessions.

The results of our second out-of-sample exercise (for the time period 2005:1-2014:4) in the Table 7 suggest that the out-of-sample forecasting performance of EPU cannot be outperformed by those of NVIX and GPR. While the pseudo R^2 values for GPR are still close to zero, and thus exhibit little explanatory power in predicting US recessions, we observe a considerable increase in the pseudo R^2 values for EPU and NVIX. For the multi-factor model, we observe that the use of all three predictors entails a deterioration of the pseudo R^2 values, which thus have less predictive power than EPU or NVIX but still better than the predictive performance of GPR. We can confirm the results obtained by Karnizova and Li (2014) and demonstrate that the EPU can be used in predicting US recessions. However, we note that the predictive power of EPU for the pre-1985 out-of-sample US recessions is very low.

6 Conclusion

A large number of models have been employed in the literature to analyze and forecast changes in output growth. The objective of this paper is to compare the forecasting ability of 10 (both uni- and bivariate) models in terms of forecasting US GNP growth at different forecasting horizons, with the bivariate models containing information on a measure of economic uncertainty. We evaluate the forecasting performance of these 10 models over the quarterly period 1919:2 to 2014:4. Our results indicate that the economic policy uncertainty index can help improve the accuracy of US GNP growth forecasts in bivariate models significantly.

Using RMSE and LPDS as model selection criteria, as well as using the SPA test and the equal predictive accuracy test, we find that the Markov Switching-Time Varying Parameter VAR (MSTVPVAR) model and the Bayesian VAR with stochastic volatility (BVARCSV) model provide accurate US GNP growth forecasts that cannot be outperformed by the other competing models used in this study. More importantly, our results show that the economic policy uncertainty measures can help improve the accuracy of US GNP growth forecasts. Furthermore, our results show that EPU and the term spread exhibit similar out-of-sample forecasting performance. It seems that other uncertainty measures such as NVIX and GPR in most cases have less explanatory power in predicting US GNP growth or recessions than EPU.

References

Adolfson, M., J. Linde, and M. Villani (2007). Forecasting performance of an open economy DSGE model. *Econometric Reviews 26*, 289–328.

Alessandri, P. and H. Mumtaz (2014). Financial regimes and uncertainty shocks. School of Economics and Finance, Queen Mary University of London Working Paper No. 729.

Avdjiev, S. and Z. Zeng (2014). Credit growth, monetary policy and economic activity in a three-regime TVAR model. *Applied Economics 46*, 2936–2951.

Bachmann, R. and C. Bayer (2011). Uncertainty business cycles-really? National Bureau of Economic Research Working Paper No. w16862.

Bachmann, R., S. Elstner, and E. Sims (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics 5*, 217–249.

Baker, S., N. Bloom, and S. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*, 1593–1636.

Balcilar, M., R. Gupta, and M. Segnon (2016). The role of economic policy uncertainty in predicting U.S. recessions: A mixed-frequency Markov-switching vector autoregressive approach. *Economics 10*, 1–21.

Bali, T. G., S. J. Brown, and Y. Tang (2015). Macroeconomic uncertainty and expected stock returns. Georgetown McDonough School of Business Research Paper No. 2407279.

Barnett, A., H. Mumtaz, and K. Theodoridis (2014). Forecasting UK GDP growth and inflation under structural change. a comparison of models with time-varying parameters. *International Journal of Forecasting 30*, 129–143.

Bekiros, S. D. and A. Paccagnini (2013). On the predictability of time-varying VAR and DSGE models. *Empirical Economics 45*, 635–664.

Bekiros, S. D. and A. Paccagnini (2014). Bayesian forecasting with small and medium scale factor-augmented vector autoregressive DSGE model. *Computational Statistics and Data Analysis 71*, 298–323.

Bekiros, S. D. and A. Paccagnini (2015a). Estimating point and density forecasts for the US economy with a factor-augmented vector autoregressive DSGE model. *Studies in Nonlinear Dynamics and Econometrics 19*, 1–30.

Bekiros, S. D. and A. Paccagnini (2015b). Macroprudential policy and forecasting using hybrid DSGE models with financial frictions and state space Markov -switching TVP-VARs. *Macroeconomic Dynamics 19*, 1565–1592.

Benati, L. (2013). Economic policy uncertainty and the great recession. Mimeo, University of Bern.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics 98*, 85–106.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*, 623–685.

Brogaard, J. and A. Detzel (2015). The asset-pricing implications of government economic policy uncertainty. *Management Science 61*, 3–18.

Caggiano, G., E. Castelnuovo, and N. Groshenny (2014). Uncertainty shocks and unemployment dynamics in US recessions. *Journal of Monetary Economics 67*, 78–92.

Caggiano, G., E. Castelnuovo, and G. Nodari (2014). Uncertainty and monetary policy in good and bad times. Dipartimento di Scienze Economiche "Marco Fanno", Working Paper No. 0188.

Campbell, J. Y. (2008). Viewpoint: Estimating the equity premium. *Canadian Journal of Economics 41*, 1–21.

Carriero, A., T. E. Clark, and M. Marcellino (2015). Common drifting volatility in large bayesian VARs. *Journal of Business and Economic Statistics*.

Castelnuovo, E., G. Caggiano, and G. Pellegrino (2015). Estimating the real effects of uncertainty shocks at the zero lower bound. Dipartimento di Scienze Economiche "Marco Fanno", Working Paper No. 0200.

Chauvet, M. and S. Potter (2013). Forecasting output. In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2. Amsterdam Netherlands: Elsevier.

Cheng, C.-H. J., W. A. Hankins, and C.-W. J. Chiu (2016). Does US partisan conflict matter for the Euro area? *Economics Letters 138*, 64–67.

Chuliá, H. S., M. Guillén, and J. M. Uribe (2015). Measuring uncertainty in the stock market. IREA-Working Paper No. IR15/024.

Clark, T. (2011). Real-time density forecasts from BVARs with stochastic volatility. *Journal of Business and Economic Statistics 29*, 327–341.

Cogley, T. and T. J. Sargent (2002). Evolving post-world war II US inflation dynamics. *NBER Macroeconomics Annual 2001 16*, 331–388.

Colombo, V. (2013). Economic policy uncertainty in the US: Does it matter for the Euro area? *Economics Letters 121*, 39–42.

Creal, D. D. and C. Wu (2017). Monetary policy uncertainty and economic fluctuations. *International Economic Review*. Forthcoming.

D'Agostino, A., L. Gambetti, and D. Giannone (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics 28*, 82–101.

Del Negro, M., R. B. Hasegawa, and F. Schorfheide (2016). Dynamic prediction pools: an investigation of financial frictions and forecasting performance. *Journal of Econometrics 192*, 391–405.

Del Negro, M. and F. Schorfheide (2011). Bayesian macroeconometrics. In J. Geweke, G. Koop, and H. V. Dijk (Eds.), *The Oxford Handbook of Bayesian Econometrics*, pp. 293–389. Oxford, University Press.

Del Negro, M. and F. Schorfheide (2012). DSGE model-based forecasting. Available at SSRN 2018451.

Diebold, F. X. and R. Mariano (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics 13*, 253–263.

Dixit, A. K. and R. S. Pindyck (1994). *Investment under uncertainty*. Princeton University Press.

Eickmeier, S., W. Lemke, and M. Marcellino (2011). Classical time-varying FAVAR models-estimation, forecasting and structural analysis. CEPR Discussion Paper No. DP8321.

Estrella, A. (1998). A new measure of fit for equations with dichotomous dependent variables. *Journal of Business & Economic Statistics 16*, 198–205.

Estrella, A. and F. Mishkin (1998). Predicting US recessions: financial variables as leading indicators. *Review of Economic and Statistics 80*, 45–61.

Giannone, D., M. Lenza, and G. Primiceri (2015). Prior selection for vector autoregressions. *Review of Economics and Statistics 97*, 436–451.

Gordon, R. J. (1986). *The American Business Cycle: Continuity and Change*. University of Chicago Press.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business and Economic Statistics 23*, 365–380.

Herbst, E. and F. Schorfheide (2012). Evaluating DSGE model forecasts of comovements. *Journal of Econometrics 171*, 152–166.

Jones, P. M. and E. Olson (2013). The time-varying correlation between uncertainty, output and inflation: Evidence from a DCC-GARCH model. *Economics Letters 118*, 33–37.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *The American Economic Review 105*, 1177–1216.

Kang, W., K. Lee, and R. A. Ratti (2014). Economic policy uncertainty and firm-level investment. *Journal of Macroeconomics 39*, 42–53.

Karnizova, L. and J. C. Li (2014). Economic policy uncertainty, financial markets and probability of US recessions. *Economics Letters 125*, 261–265.

Knotek II, E. S. and S. Khan (2011). How do households respond to uncertainty shocks. *Federal Reserve Bank of Kansas City Economic Review 96*, 5–34.

Koop, G. and D. Korobilis (2010). Bayesian multivariate time series methods. *Foundations and Trends in Econometrics 3*, 267–358.

Ludvigson, S. C., S. Ma, and S. Ng (2015). Uncertainty and business cycles: Exogenous impulse or endogenous response? National Bureau of Economic Research, Working Paper No. w21803.

Mumtaz, H. and P. Surico (2013). Policy uncertainty and aggregate fluctuations. Queen Mary University of London, School of Economics and Finance, Working Paper No. 708.

Mumtaz, H. and K. Theodoridis (2015). Common and country specific economic uncertainty. Queen Mary University of London, School of Economics and Finance, Working Paper No. 752.

Mumtaz, H. and K. Theodoridis (2016). The changing transmission of uncertainty shocks in the us: An empirical analysis. *Journal of Business and Economic Statistics*.

Mumtaz, H. and F. Zanetti (2013). The impact of the volatility of monetary policy shocks. *Journal of Money, Credit and Banking 45*, 535–558.

Nakajima, J. (2011). Time-varying parameter VAR model with stochastic volatility: An overview of methodology and empirical applications. *Monetary and Economic Studies*, 1–36.

Omay, T., R. Gupta, and G. Bonaccolto (2016). The US real GNP is trend-stationary after all. *Applied Economics Letters*, 1–5.

Primiceri, G. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies 72*, 821–852.

Rossi, B. and T. Sekhposyan (2010). Have economic models' forecasting performance for us output growth and inflation changed over time, and when? *International Journal of Forecasting 26*, 808–835.

Rossi, B. and T. Sekhposyan (2014). Evaluating predictive densities of us output growth and inflation in a large macroeconomic data set. *International Journal of Forecasting 30*, 662–682.

Rossi, B. and T. Sekhposyan (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *The American Economic Review 105*, 650–655.

Rossi, B., T. Sekhposyan, and M. Soupre (2016). Understanding the sources of macroeconomic uncertainty. Mimeo, Universitat Pompeu Fabra - Centre de Recerca en Economia Internacional (CREI).

Schorfheide, F. and D. Song (2015). Real-time forecasting with a Mixed-Frequency VAR. *Journal of Business and Economic Statistics 33*, 366–380.

Schumacher, C. (2011). Forecasting with factor models estimated on large datasets: A review of the recent literature and evidence for German gdp. *Journal of Economics and Statistics 231*, 28–49.

Shin, M. and M. Zhong (2016). A new approach to identifying the real effects of uncertainty shocks. Finance and Economics Discussion Series 2016-040. Washington: Board of Governors of the Federal Reserve System.

Stock, J. H. and M. W. Watson (2007). Why has US inflation become harder to forecast. *Journal of Money, Credit and Banking 39*, 3–33.

Wheelock, D. C. and M. E. Wohar (2009). Can the term spread predict output growth and recessions? a survey of the literature. *Federal Reserve Bank St. Review 91*, 419–440.

White, H. (2000). A reality check for data snooping. *Econometrica 68*, 1097–1126.

Wolters, J., T. Teräsvirta, and H. Lütkepohl (1998). Modeling the demand for M3 in the unified germany. *Review of Economics and Statistics 80*, 399–409.

Table 1: Descriptive statistics for US GNP growth and logarithmized EPU

Variables	Min	Max	Std	Skewness	Kurtosis	ADF	ADF*
US GNP growth	-15.458	13.518	2.966	-0.941	9.365	-13.003	-12.989
US EPU	3.631	5.666	0.401	-0.362	2.629	-5.637	-5.975

Note: ADF and ADF* represent the augmented Dickey-Fuller t-statistics of the lagged dependent variable in a regression with intercept and time trend and intercept only, respectively. The critical values for ADF and ADF∗ are **-2.869** and **-3.421**.

Table 2: RMSE for uni- and bivariate models

Models	Forecasting horizons							
	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
EPU								
ARMA	2.601	2.909	2.926	2.969	2.961	2.950	2.939	2.986
UCSV	2.863	3.462	3.421	3.657	3.878	3.786	3.750	3.853
BVAR	2.612	2.884	2.806	2.850	2.914	2.893	2.895	2.895
BVARCSV	2.583	2.852	2.793	2.892	2.943	2.923	2.931	2.942
TVAR	2.632	2.861	2.833	2.893	2.921	2.941	2.935	2.913
STVAR	2.668	7.711	35.826	65.335	128.776	189.233	246.049	301.996
MSVAR	2.608	2.884	2.830	2.883	2.925	2.899	2.909	2.910
TVPVAR	2.749	3.106	3.033	132.711	13.630	60.564	84.944	195.872
MS-TVPVAR	**2.184**	**2.486**	**2.397**	**2.595**	**2.734**	**2.569**	**2.568**	**2.607**
MFVAR	3.088	2.908	2.906	2.916	2.930	2.927	2.936	2.926
Term spread								
BVAR	2.650	2.972	2.900	3.004	3.025	2.919	2.887	2.893
BVARCSV	2.570	2.863	2.780	2.870	2.951	2.895	2.855	2.908
TVAR	2.720	3.065	2.951	3.134	3.129	2.975	2.935	2.976
STVAR	2.817	4.618	6.900	18.961	35.665	55.857	74.216	89.021
MSVAR	2.712	3.008	2.947	3.108	3.101	2.969	2.954	2.979
TVPVAR	2.777	3.112	3.104	3.101	3.482	4.913	429.660	446.085
MS-TVPVAR	2.794	**2.962**	**2.752**	**2.842**	**2.871**	**2.605**	**2.929**	**2.730**
MFVAR	3.063	3.032	2.978	2.962	2.966	2.930	2.905	2.908
NVIX								
MS-TVPVAR	2.988	3.264	3.360	3.749	3.667	3.143	3.236	3.453
GPR								
MS-TVPVAR	5.338	5.188	5.144	5.145	5.393	5.056	4.888	5.907

Note: The entries are RMSEs. The values in bold correspond to the smallest RMSEs. The EPU and term spread are used as predictors for the bivariate models to produce out-of-sample US GNP growth for 1919:2-2014:4.

Table 3: Density forecasts for uni- and bivariate models

Models	Forecasting horizons							
	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
EPU								
ARMA	4.125	4.302	4.464	4.614	4.842	5.123	5.547	6.091
UCSV	4.481	4.785	4.979	5.141	5.303	5.423	5.553	5.688
BVAR	**3.942**	4.084	4.161	4.210	4.228	4.228	4.236	**4.234**
BVARCSV	3.979	4.071	**4.080**	**4.091**	**4.199**	**4.046**	**4.230**	4.271
TVAR	3.945	4.120	4.138	4.216	4.248	4.261	4.270	4.276
STVAR	3.989	4.461	4.948	5.405	5.778	5.815	6.197	6.382
MSVAR	4.107	**4.063**	4.204	4.228	4.260	4.263	4.295	4.298
TVPVAR	3.985	4.778	5.240	5.561	5.858	6.048	6.150	6.268
MS-TVPVAR	5.165	6.294	7.492	7.224	6.976	8.975	7.647	7.576
MFVAR	4.255	4.204	4.220	4.231	4.246	4.260	4.266	4.257
Term spread								
BVAR	3.969	4.152	4.230	4.270	4.290	4.271	4.255	4.252
BVARCSV	3.877	**4.025**	**4.058**	**4.046**	**4.244**	4.394	**4.012**	**4.142**
TVAR	3.947	4.156	4.209	4.315	4.347	4.350	4.338	4.350
STVAR	4.013	4.645	5.272	5.719	6.118	6.581	6.822	7.100
MSVAR	**3.941**	4.161	4.286	4.310	4.380	4.347	4.341	4.347
TVPVAR	4.004	4.720	5.070	5.286	5.524	5.647	5.696	5.854
MS-TVPVAR	4.811	5.472	6.032	5.872	6.223	6.979	6.273	6.050
MFVAR	4.333	4.294	4.311	4.308	4.341	4.328	4.291	4.289
NVIX								
BVARCSV	**3.812**	4.137	4.200	4.215	4.251	**4.234**	4.101	4.279
GPR								
BVARCSV	3.895	4.105	4.197	4.237	4.309	4.393	4.312	4.341

Note: The entries are log predictive density scores (LPDS). The values in bold correspond to the smallest LPDS. The EPU and term spread are used as predictors for the bivariate models to produce out-of-sample US GNP growth for 1919:2-2014:4.

Table 4: Results of SPA tests for US-GNP growth forecasts

Basic model	Forecasting horizons							
	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
EPU								
ARMA	**0.065**	0.223	0.132	0.202	0.508	0.332	0.323	0.228
UCSV	**0.045**	**0.008**	**0.018**	**0.008**	**0.004**	**0.007**	**0.010**	**0.009**
BVAR	0.112	0.293	0.209	0.673	0.791	0.499	0.386	0.416
BVARCSV	0.120	0.291	0.194	0.412	0.524	0.316	0.302	0.354
TVAR	0.079	0.311	0.215	0.446	0.573	0.305	0.288	0.401
STVAR	**0.081**	**0.059**	**0.091**	**0.094**	**0.096**	**0.098**	**0.083**	**0.081**
MSVAR	0.116	0.325	**0.086**	**0.044**	0.695	0.513	0.404	0.313
TVPVAR	**0.035**	**0.050**	**0.074**	0.253	0.116	0.177	0.141	0.200
MSTVPVAR	1.000	0.894	1.000	0.962	0.812	0.913	0.980	0.988
MFVAR	**0.008**	0.202	0.117	0.204	0.639	0.371	0.174	0.364
Term spread								
ARMA	0.708	0.721	0.136	0.212	0.763	0.337	0.197	0.247
UCSV	**0.042**	**0.007**	**0.016**	**0.005**	**0.003**	**0.006**	**0.009**	**0.007**
BVAR	0.175	0.293	0.267	0.215	0.109	0.444	0.872	0.785
BVARCSV	0.815	0.924	0.447	0.717	0.739	0.392	0.978	0.721
TVAR	0.146	0.112	0.106	0.037	0.133	0.241	0.330	0.245
STVAR	**0.077**	0.036	**0.079**	**0.081**	**0.082**	**0.082**	**0.050**	**0.010**
MSVAR	**0.032**	0.109	0.171	0.031	0.075	0.326	0.159	**0.083**
TVPVAR	**0.049**	**0.063**	**0.004**	**0.042**	0.126	**0.098**	0.141	0.135
MSTVPVAR	0.357	0.520	0.676	0.664	0.721	0.974	0.666	0.948
MFVAR	**0.034**	0.291	0.122	0.429	0.624	0.437	0.759	0.761

The numbers in the Table represent the p-values of the SPA test of Hansen (2005) for the pertinent model and criterion. We test the null hypothesis that a benchmark model outperforms other candidate models. The p-values of the SPA test that are smaller than or equal to the 10% confidence level under a pre-specified loss function are in bold. The EPU and term spread are used as predictors for the bivariate models to produce out-of-sample US GNP growth for 1919:2-2014:4.

Table 5: Results of equal accuracy tests based on the point forecast measure for US-GNP growth forecasts

Model1	model2	Forecasting horizons							
		1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
		EPU							
BVAR	MSTVPVAR	1.979	1.401	1.278	0.774	0.632	1.116	0.930	0.8468
		(0.024)	(0.081)	(0.101)	(0.220)	(0.264)	(0.132)	(0.176)	(0.199)
BVARCSV		1.826	1.196	1.317	0.977	0.869	1.473	1.165	1.138
		(0.034)	(0.116)	(0.094)	(0.165)	(0.193)	(0.071)	(0.122)	(0.128)
TVAR		2.088	1.276	1.344	0.927	0.656	1.261	1.030	0.890
		(0.019)	(0.101)	(0.090)	(0.177)	(0.256)	(0.104)	(0.152)	(0.187)
STVAR		1.767	1.801	1.239	1.168	1.178	1.243	1.324	1.080
		(0.039)	(0.036)	(0.108)	(0.122)	(0.120)	(0.107)	(0.093)	(0.140)
MSVAR		1.983	1.394	1.349	0.875	0.676	1.134	0.958	0.888
		(0.024)	(0.082)	(0.089)	(0.191)	(0.250)	(0.129)	(0.169)	(0.187)
TVPVAR		2.422	1.942	1.728	1.000	1.018	0.999	0.995	0.998
		(0.008)	(0.026)	(0.042)	(0.159)	(0.155)	(0.159)	(0.160)	(0.160)
MFVAR		3.429	1.338	1.590	0.993	0.691	1.259	1.052	0.948
		(0.000)	(0.091)	(0.056)	(0.161)	(0.245)	(0.104)	(0.147)	(0.172)
		Term spread							
BVAR	MSTVPVAR	-0.565	0.039	0.465	0.518	0.520	0.998	-0.141	0.594
		(0.714)	(0.485)	(0.321)	(0.302)	(0.302)	(0.160)	(0.556)	(0.276)
BVARCSV		-0.876	-0.448	0.107	0.113	0.355	1.032	-0.276	0.694
		(0.809)	(0.673)	(0.457)	(0.455)	(0.361)	(0.151)	(0.609)	(0.244)
TVAR		-0.255	0.339	0.634	0.955	0.943	1.149	0.024	0.877
		(0.601)	(0.367)	(0.263)	(0.170)	(0.173)	(0.126)	(0.491)	(0.191)
STVAR		0.074	2.002	1.638	1.343	1.381	1.482	1.815	2.240
		(0.471)	(0.023)	(0.051)	(0.090)	(0.084)	(0.070)	(0.035)	(0.013)
MSVAR		-0.315	0.183	0.639	0.914	0.762	1.131	0.085	0.878
		(0.623)	(0.427)	(0.262)	(0.181)	(0.223)	(0.129)	(0.466)	(0.190)
TVPVAR		-0.062	0.642	1.196	0.867	1.474	1.193	1.000	0.998
		(0.525)	(0.261)	(0.116)	(0.193)	(0.071)	(0.117)	(0.159)	(0.160)
MFVAR		1.184	0.257	0.690	0.383	0.299	1.017	-0.077	0.610
		(0.118)	(0.398)	(0.245)	(0.351)	(0.382)	(0.155)	(0.531)	(0.271)

Note: Table entries are t-statistics and p-values (in parentheses) of the Diebold and Mariano (1995) test. We test the null hypothesis that the forecasts at horizon *h* of model1 are equal to that of model2 against the one-sided alternative that forecasts from model1 is inferior to those of model 2. The EPU and term spread are used as predictor the bivariate models to produce out-of-sample US GNP growth for 1919:2-2014:4.

Table 6: Results of equal accuracy tests based on density forecast measure for US-GNP growth forecasts

Model1	model2	Forecasting horizons							
		1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
		EPU							
BVAR	BVARCSV	-0.256	0.081	0.522	0.513	0.105	0.795	0.021	-0.101
		(0.601)	(0.468)	(0.301)	(0.304)	(0.458)	(0.213)	(0.491)	(0.540)
TVAR		-0.224	0.255	0.423	0.594	0.184	0.958	0.133	0.016
		(0.589)	(0.400)	(0.336)	(0.276)	(0.427)	(0.169)	(0.447)	(0.494)
STVAR		0.074	1.978	3.911	4.284	4.789	5.503	4.620	4.148
		(0.471)	(0.024)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MSVAR		0.470	-0.038	0.531	0.446	0.177	0.702	0.178	0.066
		(0.319)	(0.515)	(0.298)	(0.328)	(0.430)	(0.242)	(0.430)	(0.474)
TVPVAR		0.060	3.960	5.374	5.312	4.594	5.078	3.978	3.725
		(0.487)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MSTVPVAR		4.463	4.074	3.925	3.671	3.327	3.062	3.470	3.085
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)
MFVAR		1.802	0.870	1.072	0.823	0.202	1.168	0.129	-0.040
		(0.036)	(0.193)	(0.142)	(0.205)	(0.420)	(0.122)	(0.449)	(0.516)
		Term spread							
BVAR	BVARCSV	0.487	0.624	0.790	1.071	0.144	-0.308	1.187	0.374
		(0.313)	(0.266)	(0.215)	(0.143)	(0.443)	(0.621)	(0.118)	(0.354)
TVAR		0.393	0.685	0.802	1.362	0.338	-0.111	1.550	0.713
		(0.347)	(0.247)	(0.212)	(0.087)	(0.368)	(0.544)	(0.061)	(0.238)
STVAR		0.712	2.713	4.238	5.288	4.523	4.375	5.106	4.844
		(0.238)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
MSVAR		0.312	0.526	0.817	0.915	0.364	-0.105	1.141	0.573
		(0.378)	(0.300)	(0.207)	(0.180)	(0.358)	(0.542)	(0.127)	(0.283)
TVPVAR		0.786	3.139	3.939	4.804	3.327	2.645	4.434	3.858
		(0.216)	(0.001)	(0.000)	(0.000)	(0.000)	(0.004)	(0.000)	(0.0001)
MSTVPVAR		4.649	4.977	4.269	4.064	3.708	3.232	4.414	4.186
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
MFVAR		2.479	1.443	1.406	1.649	0.329	-0.172	1.632	0.555
		(0.007)	(0.075)	(0.080)	(0.050)	(0.371)	(0.568)	(0.052)	(0.290)

Note: Table entries are t-statistics and p-values (in parentheses) of the Diebold and Mariano (1995) test. We test the null hypothesis that the forecasts at horizon *h* of model1 are equal to those of model2 against the one-sided alternative that forecasts from model1 is inferior to that of model 2. The EPU and term spread are used as predictors for the bivariate models to produce out-of-sample US GNP growth for 1919:2-2014:4.

Table 7: Out-of-sample pseudo R^2's

Variables (x_t)	Forecasting horizons							
$Pr(R_{t+h} = 1) = \Phi(\varsigma + \beta x_t)$								
	1Q	2Q	3Q	4Q	5Q	6Q	7Q	8Q
1985Q1 - 2014Q3								
EPU	0.166	0.168	0.170	0.172	0.175	0.178	0.181	0.183
1985Q1 - 2014Q4								
NVIX	0.167	0.165	0.163	0.161	0.159	0.157	0.155	0.153
GPR	0.018	0.019	0.021	0.024	0.029	0.034	0.040	0.048
1985Q1 - 2014Q3								
EPU + NVIX + GPR	0.154	0.153	0.153	0.154	0.156	0.159	0.162	0.166
1900Q1 - 2014Q3								
EPU	0.019	0.019	0.019	0.019	0.019	0.018	0.018	0.018
1900Q1 - 2014Q4								
NVIX	0.019	0.019	0.018	0.018	0.018	0.018	0.018	0.018
GPR	0.055	0.056	0.056	0.057	0.057	0.057	0.058	0.059
1900Q1 - 2014Q3								
EPU + NVIX + GPR	0.067	0.067	0.068	0.068	0.069	0.069	0.069	0.070

Note: The pseudo R^2 is computed as defined in Estrella (1998). A value of the pseudo R^2 that approaches zero indicates poor fit, while a value that approaches 1 indicates a good fit.
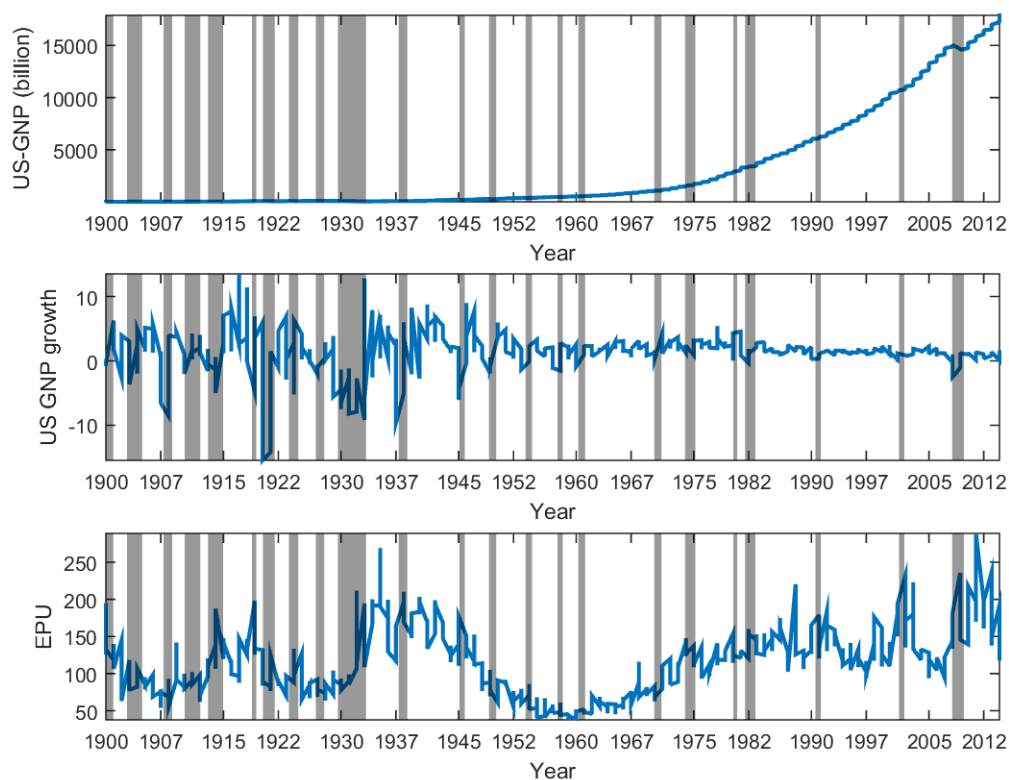
Figure 1: Plot of quarterly US real GNP, growth rates and the economic policy uncertainty index for the period 1900:1-2014:4. Vertical bars (gray) represent NBER-dated recessions.
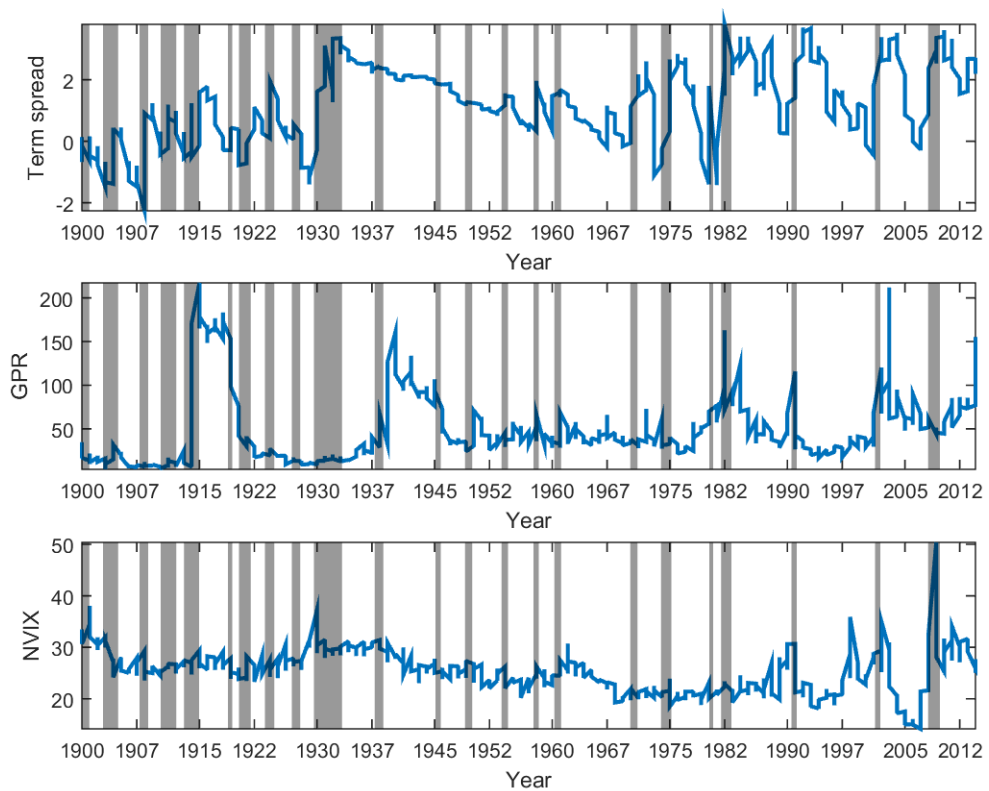
Figure 2: Plot of quarterly US term spread, geopolitical risk and volatility index for the period 1900:1-2014:4. Vertical bars (gray) are NBER-dated recessions.
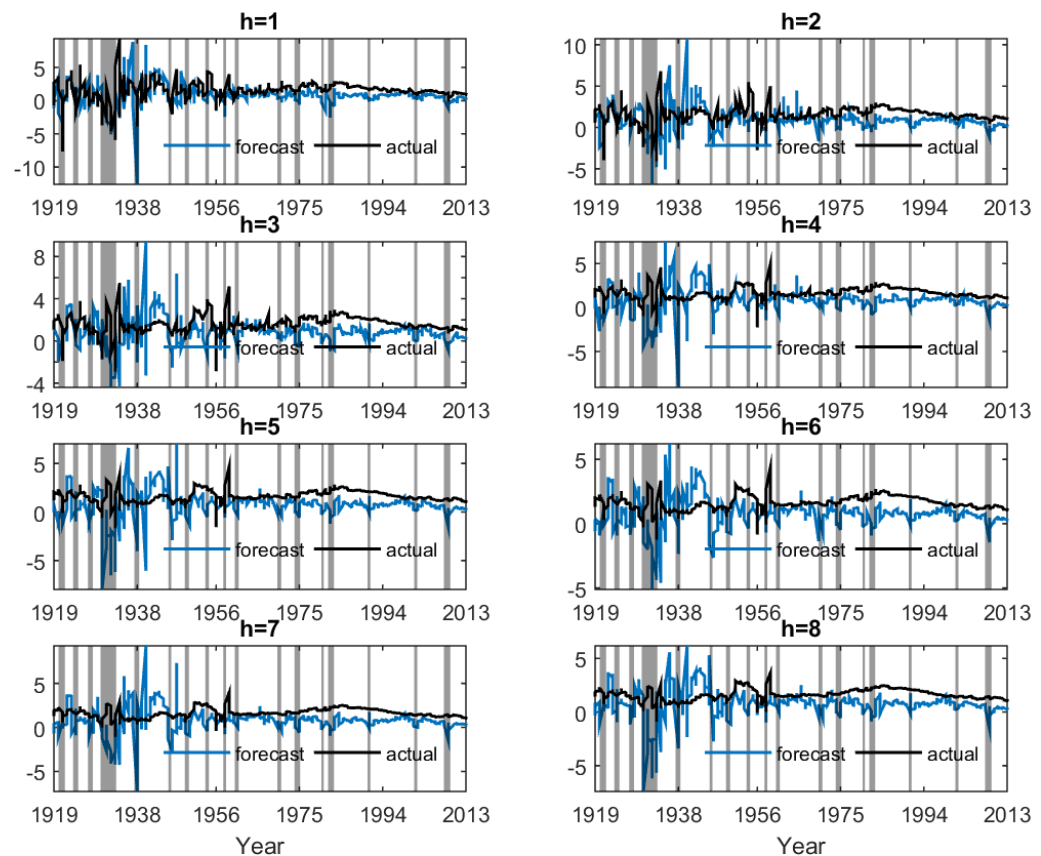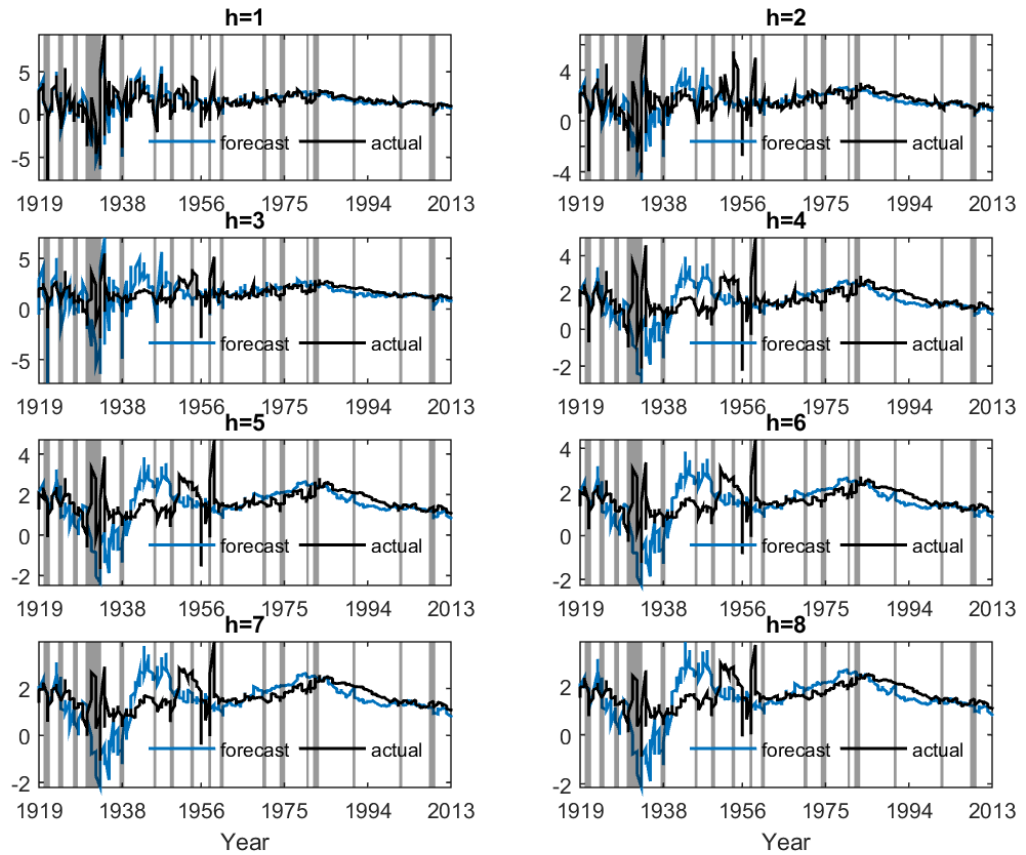
Figure 3: Plot of quarterly US GNP growth out-of-sample forecasts obtained from the MSTVPVAR model, using the EPU for 1919:2 to 2014:4 for different forecasting horizons (1- up to 8-quarter ahead) and actual real US GNP growth. NBER-dated recessions are represented by the vertical bars (gray).
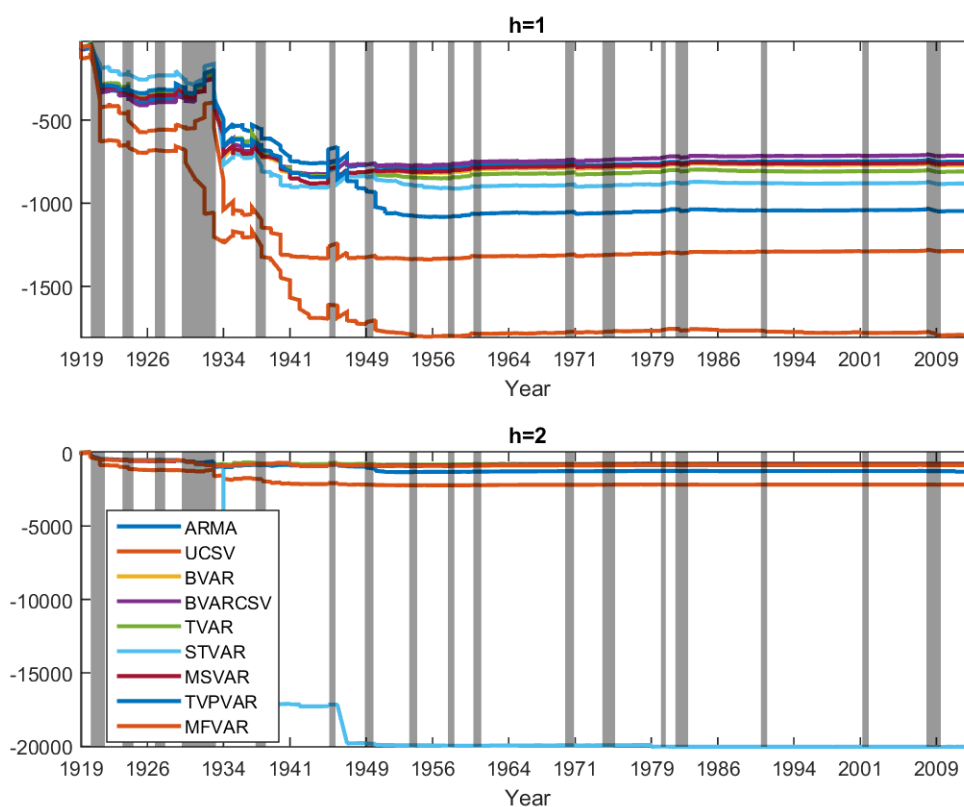
Figure 4: Plot of quarterly US GNP growth out-of-sample forecasts obtained from the BVAR model with stochastic volatility, using the EPU for 1919:2-2014:4 for different forecasting horizons (1- up to 8-quarter ahead) and actual real US GNP growth. NBER-dated recessions are represented by the vertical bars (gray).

Figure 5: Plot of cumulative differences in density forecast errors for quarterly US GNP
growth out-of-sample forecasts for MSTVPVAR relative to competitors, using
EPU for the period 1919:2-2014:4 at forecasting horizons 1- and 2-quarter
ahead. NBER-dated recessions are presented in vertical bars (gray).
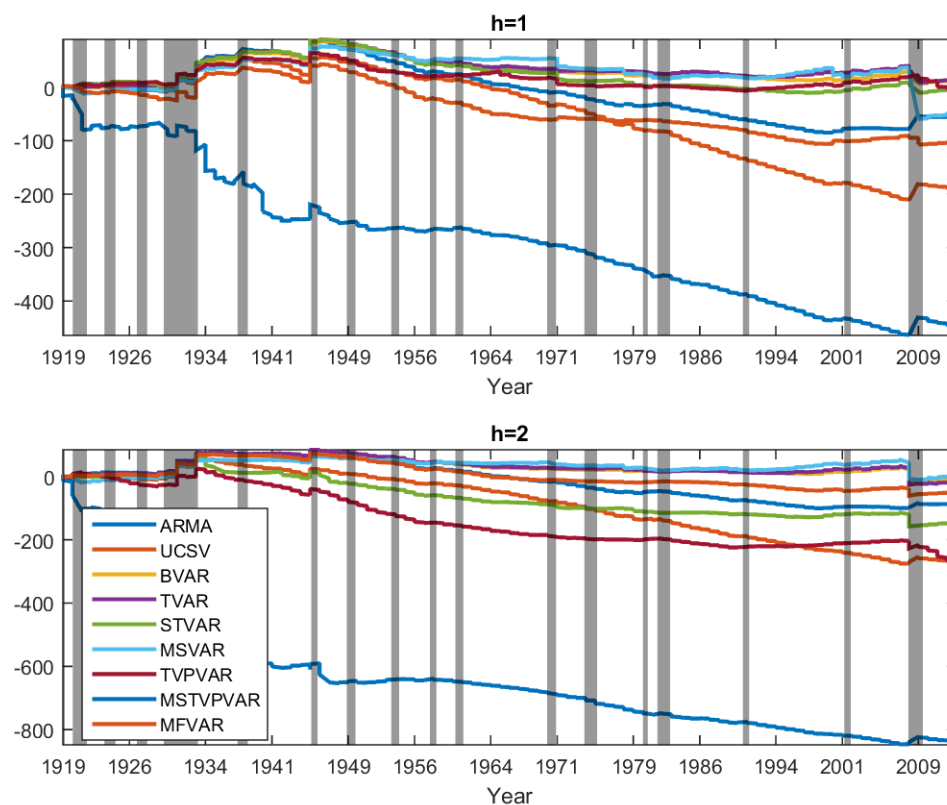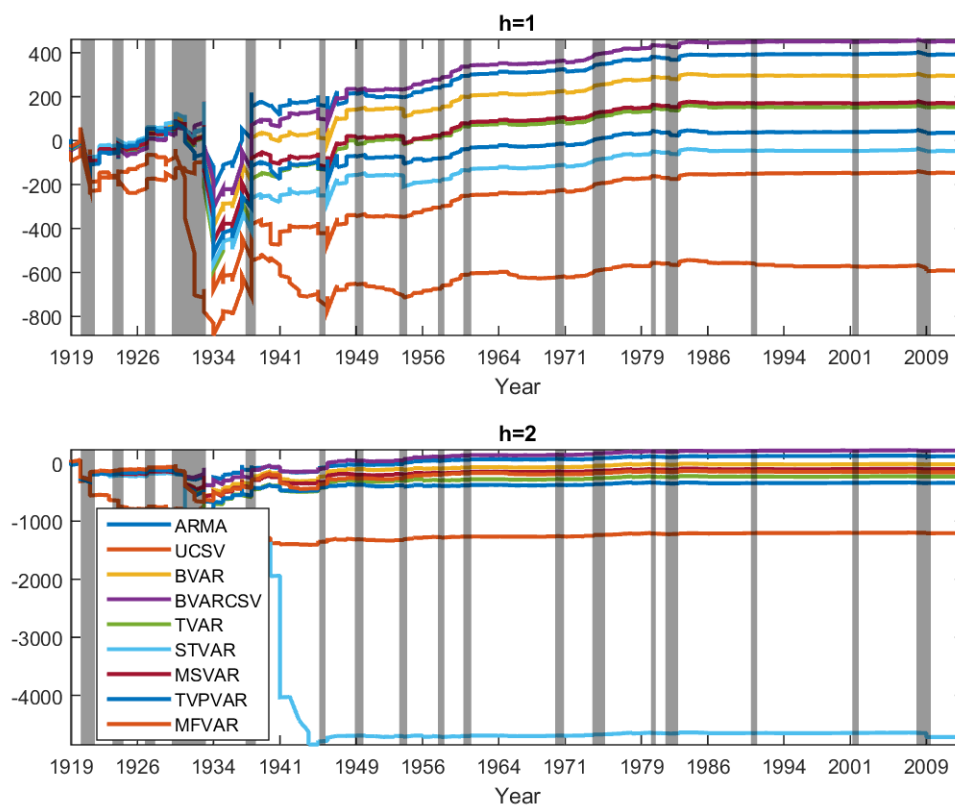
Figure 6: Plot of cumulative differences in density forecast errors for quarterly US GNP growth out-of-sample forecasts for BVARCSV relative to competitors, using EPU for the period 1919:2-2014:4 at forecasting horizons 1- and 2-quarter ahead. NBER-dated recessions are presented in vertical bars (gray).

Figure 7: Plot of cumulative differences in density forecast errors for quarterly US GNP growth out-of-sample forecasts for MSTVPVAR relative to competitors, using the term spread for the period 1919:2-2014:4 at forecasting horizons 1- and 2-quarter ahead. NBER-dated recessions are presented in vertical bars (gray).
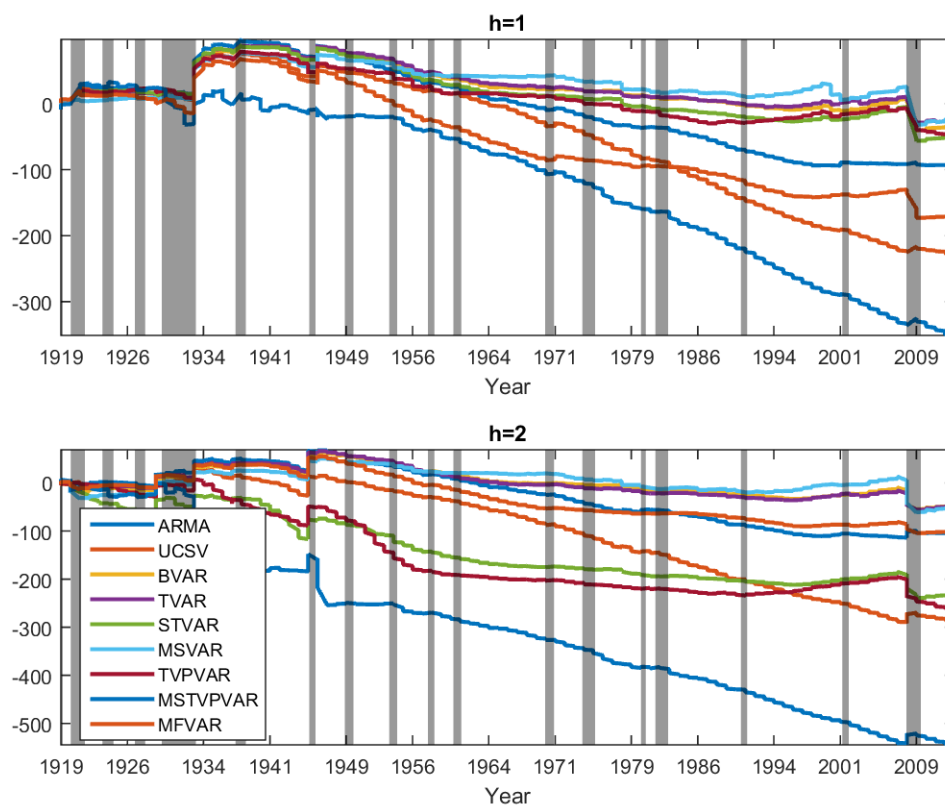
Figure 8: Plot of cumulative differences in density forecast errors for quarterly US GNP growth out-of-sample forecasts for BVARCSV relative to competitors, using the term spread for the period 1919:2-2014:4 at forecasting horizons 1- and 2-quarter ahead. NBER-dated recessions are presented in vertical bars (gray).