Received: 27 April 2020 | Revised: 17 June 2022 | Accepted: 30 June 2022

DOI: 10.1002/jae.2944

WILEY Journal of APPLIED ECONOMETRICS

RESEARCH ARTICLE

Density forecasting with Bayesian Vector Autoregressive models under macroeconomic data uncertainty

Michael P. Clements[1] | **Ana Beatriz Galvão**[2]

[1]ICMA Centre, Henley Business School, University of Reading, Reading, UK

[2]Warwick Business School, University of Warwick, Warwick, UK

Correspondence
Michael P. Clements, ICMA Centre, Henley Business School, University of Reading, Reading, UK.
Email: m.p.clements@reading.ac.uk

Summary

Macroeconomic data are subject to data revisions. Yet, the usual way of generating real-time density forecasts from Bayesian Vector Autoregressive (BVAR) models makes no allowance for data uncertainty from future data revisions. We develop methods of allowing for data uncertainty when forecasting with BVAR models with stochastic volatility. First, the BVAR forecasting model is estimated on real-time vintages. Second, the BVAR model is jointly estimated with a model of data revisions such that forecasts are conditioned on estimates of the 'true' values. We find that this second method generally improves upon conventional practice for density forecasting, especially for the United States.

KEYWORDS

inflation and output growth predictive densities, real-time forecasting, real-time-vintages, stochastic volatility

1 | INTRODUCTION

Decision makers employ probabilistic forecasts of macroeconomic variables to compute the probability of future outcomes of interest as an aid to determining which course of action to take. For example, based on density forecasts, one can quantify the probability of sluggish growth (say lower than 1%) and/or of deflation to support a monetary policy decision. This paper considers whether it is important to make an allowance for data uncertainty when computing probabilistic forecasts in real-time, given that most macroeconomic variables are subject to data revisions.

The literature addresses this question for models that assume the variances of the disturbances are homoscedastic but does not consider time-varying volatility. Clements (2017) shows that the standard real-time approach, which estimates the forecasting model on the vintage of data available at the forecast origin, will likely give an inaccurate assessment of the uncertainty surrounding future values of the variables, especially of the early-vintage estimates of those values. He considers autoregressive models with constant-variance disturbances, and his work follows on from the work on point forecasts of Koenig et al. (2003) and Clements and Galvão (2013b), inter alia. Yet the recent literature on macroeconomic forecasting strongly supports the use of multivariate models with time-varying conditional volatility to deliver accurate density forecasts in real time (see, e.g., Carriero et al., 2020; Clark, 2011; Clark & Ravazzolo, 2014; Diebold et al., 2017). We consider the impact of data uncertainty for multivariate models with time-varying volatility.

Although the literature on forecasting with multivariate models with time-varying volatility uses real-time data, that is, the vintages of macroeconomic time series that were actually available at the time the forecast was made, it does not explicitly consider the impact of data revisions on the measurement of forecasting uncertainty. The conventional approach to real-time forecasting that underlies this literature—using the vintage of data available at the forecast origin—fails to make an allowance for data uncertainty in the estimation of the model, or in the generation of forecasts from the model. The conventional approach is sometimes known as end-of-sample, abbreviated to EOS. It regards the data as given, and does not allow for the consequences of the data being revised over time. However, making an allowance for data uncertainty in the evaluation of the forecasts from the model is commonplace: Forecasts are routinely compared with advance estimates of the actual data, or first-finals, or the latest vintage of data available when the study is undertaken.

Our main aim is to assess whether the conventional way of forecasting in real-time can be improved upon for Bayesian Vector Autoregressive (BVARs; see, e.g., Doan et al., 1984; Sims, 1980) with stochastic volatility. These are multivariate models that are popular in probabilistic macro forecasting (Clark, 2011; Clark & Ravazzolo, 2014; Carriero et al., 2020). Our first major contribution facilitates this assessment by adapting existing modelling approaches to allow for revision-driven data uncertainty. We consider two ways of allowing for revision-driven data uncertainty, and these are compared with the conventional approach of using the latest vintage of the time series available at each point in time.

The first of our two approaches is the use of real-time-vintage (RTV) data, advocated by Koenig et al. (2003) and Clements and Galvão (2013b) for point forecasting, and shown by Clements (2017) to be a simple and effective way of delivering more accurate assessments of forecasting uncertainty in univariate AR models when the error variance is homoscedastic. The second approach is based on Kishor and Koenig (2012) (KK), who propose estimating the VAR on the 'true values' of the variables at the same time as modelling the revisions.[1] The extension of this approach to allow for stochastic volatility is more involved, and we develop a Bayesian approach to accomplish this task. We detail how the Bayesian approach allows the KK approach to be applied to quarterly data with more rounds of revisions than in the application of Kishor and Koenig (2012) to annual data with few revisions, how it allows the generation of density forecasts, and how it allows for stochastic volatility. We also detail our precise empirical implementation of the extended approach, setting out how modelling choices are guided by the properties of the data revisions to the series we consider.

Our second main contribution is empirical. We undertake a thorough examination of the forecast performance of prototypical BVAR models for the United States and United Kingdom, to determine the relevance of revision-driven data uncertainty. In principle such data uncertainty should matter: We establish this in Appendix A. However, the analytical results necessarily abstract from certain aspects of the actual forecasting environment, such as parameter nonconstancies and small-sample parameter estimation uncertainty. The simplified settings we consider analytically serve to illustrate some of the factors that may shape the empirical findings, such as the properties of the revisions, but we regard it as first and foremost an empirical matter as to whether allowing for data uncertainty has a large effect on forecast accuracy. Our empirical findings are supplemented with a Monte Carlo designed to further explore the factors that determine the impact of data uncertainty on forecast performance.

Because we are interested in potential improvements from allowing for data uncertainty, our focus is on relative measures of density forecasting performance (compared with a benchmark that ignores data uncertainty), rather than absolute tests for correct specification such as the ones considered in Rossi and Sekhposyan (2013).[2] In simplified settings, we can show that RTV maximizes the out-of-sample real-time log score, and this motivates our interest in RTV as a method to deal with data uncertainty when forecasting with BVARs in real-time. Clearly, if the KK model provides an accurate description of the real-time forecasting environment, by explicitly modelling data revisions, it will correctly account for data uncertainty, and so should also provide improved density forecasts.

To anticipate our empirical findings, we find that allowing for data uncertainty using the KK approach generally improves the accuracy of density forecasts of BVAR models with stochastic volatility for variables subject to revisions such as output growth and inflation. For the United States, we generally find statistically significant improvements. For the United Kingdom, we only find significant improvements at the year ahead horizon. The RTV approach significantly improves the performance of BVAR models with *constant* volatility for these variables, but is not advantageous for models with stochastic volatility. A simulation exercise suggests that the inclusion of stochastic volatility compensates for some of the detrimental effects of data revisions on BVAR density forecasting that would arise were we to use real-time end-of-sample data rather than the RTV approach. That is, the inclusion of SV tempers the gains that would otherwise accrue to RTV.

The plan of the remainder of the paper is as follows. Section 2 describes ways of allowing for revision-driven data uncertainty, with a discussion of our Bayesian method of estimating the KK model extended to allow for stochastic uncertainty, and of the factors that guide the empirical implementation of this method. Section 3 sets out the empirical forecasting

[1] In a sense, 'true' values are never observed, and the definition is usually used pragmatically to refer to the maturity after which further revisions are unpredictable and largely inconsequential.

[2] For a survey of these techniques, see Corradi and Swanson (2006).

comparisons we undertake and considers the implications of accounting for data uncertainty for measuring the predictive variance (or forecasting uncertainty). The forecasting comparisons mainly focus on the value of allowing for data uncertainty in models with stochastic volatility. Section 4 also presents a Monte Carlo study to evaluate the expected relative performance of RTV applied to BVAR models. Section 5 offers some concluding remarks.

2 | METHODS TO DEAL WITH DATA UNCERTAINTY IN BVAR FORECASTING

Clements and Galvão (2013b) show that if real-time data are reorganized into 'real-time vintages' (RTV) for model estimation, instead of employing the conventional end-of-sample approach, the real-time accuracy of point forecasts from autoregressive models may be improved. This is one of the methods we consider here to improve real-time forecasting with BVARs by taking into account data uncertainty. The other method is based on Kishor and Koenig (2012) (KK), who propose estimating the VAR using 'true' values of the variables subject to revision, as part of a system that also includes equations to model the dynamics of data revisions. We extend KK to incorporate stochastic volatility (SV) to allow for time-variation in expected forecast uncertainty. This puts the KK approach on the same footing as the BVAR models with SV.

Below we consider the BVAR with RTV, followed by the KK approach, including the extension to allow for SV.

2.1 | BVAR with real-time vintages

Consider the simple case of a forecaster using a vector autoregressive model of order p for forecasting in real time. If the forecaster employs the latest-available vintage, that is, EOS, she will estimate the following model:

$$\mathbf{y}_t^{T+1} = \boldsymbol{\beta}^{EOS}\mathbf{x}_{t-1}^{T+1} + \mathbf{e}_t^{EOS}, \quad \text{for } t = p+1, .., T, \tag{1}$$

where $\mathbf{y}_t^{T+1}$ is a $N \times 1$ vector of the vintage $T+1$ estimate of each variable for the reference period t value, where t runs from $p+1$ up to T. The lags are also obtained from the latest vintage as $\mathbf{x}_{t-1}^{T+1} = (1, \mathbf{y}_{t-1}^{T+1'}, \ldots, \mathbf{y}_{t-p}^{T+1'})'$, implying that $\boldsymbol{\beta}^{EOS}$ is $N \times (Np+1)$ matrix. We assume the data are published with a one period delay.

If the forecaster has access to $T-1$ past vintages of the endogenous variables, that is, she has access to a real-time database, then RTV for the VAR(p) model is given by estimation of

$$\mathbf{y}_t^{t+1} = \boldsymbol{\beta}^{RTV}\mathbf{x}_{t-1}^{t} + \mathbf{e}_t^{RTV}, \text{ for } t = p+1, .., T, \tag{2}$$

where $\mathbf{y}_t^{t+1}$ is the first estimate of each endogenous variable for reference period t such that $\left\{\mathbf{y}_t^{t+1}\right\}_{t=p+1}^{t=T}$ is the time series of first releases for each variable. The $Np+1$ vector of right-hand side variables consist of

$$\mathbf{x}_{t-1}^{t} = (1, \mathbf{y}_{t-1}^{t'}, \mathbf{y}_{t-2}^{t'}, \ldots, \mathbf{y}_{t-p}^{t'})', \text{ for } t = p+1, \ldots, T;$$

that is, all lags are taken from the vintage at t, so for more than one lag ($p > 1$), at least partially revised data are used.

When forecasting in real-time, the forecasts are typically conditioned on initial and early releases, that is, on $\mathbf{x}_T^{T+1}$ (and $\mathbf{x}_{T-1}^{T+1}$ etc. depending on p).[3] As described by Clements and Galvão (2013b), estimating the VAR with RTV data gives theoretically optimal forecasts for the first-release, as opposed to using EOS. Hence the use of RTV is a promising way of dealing with variables subject to data revisions. Clements and Galvão (2013b) apply RTV to univariate models and predictive regressions, but they do not evaluate the forecasting performance of RTV applied to VAR models.

As discussed in Clements (2017) for autoregressive forecasting models, the predictive densities of EOS and RTV-estimated models will differ, and will depend on the nature of the data revisions. Based on the statistical model of data revisions described in Clements (2017), the one-step-ahead predictive variance of an AR(1) estimated with EOS data, $\sigma_{T+1|T}^{2,EOS}$, will be larger than the predictive variance of the same AR(1) model estimated with RTV data, $\sigma_{T+1|T}^{2,RTV}$, if data revisions are news, as defined by Mankiw and Shapiro (1986). News revisions are driven by the incorporation of information that was not available when the earlier estimates was made. In the case of noise, revisions embody a reduction in the earlier estimates' measurement error, and the variance of the more mature estimates will be smaller than that of the earlier

[3] By 'real time', we mean feasible forecasts that could have been made at the time the forecasts are assumed to have been made.

estimates. Clements (2017) shows that the predictive variances will satisfy $\sigma^{2,RTV}_{T+1|T} > \sigma^{2,EOS}_{T+1|T}$ for noise. Whether revisions are news or noise, if the aim is to forecast the first estimate of a macroeconomic variable, the predictive variance estimate using RTV data ought to be more accurate, and so produce more accurate density forecasts.

In the Appendix, we extend this analysis. We show that RTV will deliver better one-step-ahead density forecasting performance (as measured by the logscore) than EOS if the target is the first estimate. The expected improvement in density forecasting performance, measured by the log score, holds for news and noise revisions, and the relative improvement over EOS is increasing in the size of revision (as measured by their variance). The analysis for an AR(1) forecasting model is a motivation to consider RTV as alternative to EOS when employing VAR forecasting models for density forecasting of variables subject to data revisions.

2.1.1 | Adding stochastic volatility

In the macroeconomic forecasting literature, stochastic volatility has been found to play a key role in density forecasting (see in particular Clark & Ravazzolo, 2014). Hence, we allow for time-variation in the volatility of the disturbances in Equations (1) and (2) by allowing for a random walk process for the conditional variances. We choose the BVAR-SV specification and estimation algorithm used by Carriero et al. (2019), including the corrections described in Carriero et al. (2022).[4] The specification is such that the variance of each disturbance in the VAR may change slowly over time, but the covariances are fixed.

We apply the same estimation algorithm to obtain the SV specifications of (1) and (2).[5] Draws from the predicted density are obtained using 5000 draws from the posterior densities of the parameters (including the variance-covariance matrix of the disturbances) after the initial 5000 Gibbs draws were discarded with multistep forecasts obtained by iteration, including draws from the disturbances. We compute point forecasts and their forecasting uncertainty using the mean and the variance of the predicted density draws for each horizon.

2.2 | The KK BVAR approach

An alternative to RTV to deal with data uncertainty, when forecasting variables subject to data revisions, is the approach proposed by Kishor and Koenig (2012), henceforth KK. KK propose estimating the VAR(p) on the 'true values' of the endogenous variables, that is:

$$\mathbf{y}_t = \mathbf{c}^{KK} + \boldsymbol{\beta}^{KK}\mathbf{x}_{t-1} + \mathbf{e}^{KK}_t, \quad \text{for } t = p+1, .., T, \tag{3}$$

where $\mathbf{x}_{t-1} = (\mathbf{y}'_{t-1}, \ldots, \mathbf{y}'_{t-p})'$. The problem of course is that we do not observe true (or even the revised) values of all the observations at time $T+1$, the forecast origin. The solution is to simultaneously model data revisions to provide 'true values' up to T. The main assumption is that the true values, or an efficient estimate of these, is available l quarters after the reference quarter, that is, $\mathbf{y}_t = \mathbf{y}^{t+l}_t$ for $t = p+1, \ldots, T-l+1$. For the remaining observations, $\mathbf{y}_{T-l+2}, \ldots, \mathbf{y}_T$, KK suggest using a system of equations and the Kalman Filter.

At $t+1$, we assume we observe l estimates of $N \times 1$ vector $\mathbf{y}_t$: $\mathbf{z}^{t+1}_t = (\mathbf{y}^{t+1'}_t, \mathbf{y}^{t+1'}_{t-1}, \ldots, \mathbf{y}^{t+1'}_{t-l+1})'$ for $t = l, \ldots, T$, where the assumption that true values are observed imply that $\mathbf{y}^{t+1}_{t-l+1} = \mathbf{y}_{t-l+1}$. Let $\mathbf{rev}_t$ denote the $N(l-1) \times 1$ vector of revisions defined as the difference between the first $(l-1)$ vectors of observed values $(\mathbf{y}^{t+1'}_t, \mathbf{y}^{t+1'}_{t-1}, \ldots, \mathbf{y}^{t+1'}_{t-l+2})'$ and true values $\mathbf{y}_t$ as

$$\begin{aligned}\mathbf{rev}_t &= ((\mathbf{y}^{t+1}_t - \mathbf{y}_t)', (\mathbf{y}^{t+1}_{t-1} - \mathbf{y}_{t-1})', \ldots, (\mathbf{y}^{t+1}_{t-l+2} - \mathbf{y}_{t-l+2})')' \\ &= (\mathbf{rev}^{(1)'}_t, \mathbf{rev}^{(2)'}_{t-1}, \ldots, \mathbf{rev}^{(l-1)'}_{t-l+2})'\end{aligned} \tag{4}$$

for $l > 2$. If the second estimate reveals true values as in Kishor and Koenig (2012), $l = 2$, then, $\mathbf{rev}_t = (\mathbf{y}^{t+1}_t - \mathbf{y}_t)$.

We model the $(l-1)$ data revisions processes for each endogenous variable in $\mathbf{y}_t$ using a VAR(q):

$$\mathbf{rev}_t = \mathbf{k}_0 + \mathbf{K}_1\mathbf{rev}_{t-1} + \ldots + \mathbf{K}_q\mathbf{rev}_{t-q} + \mathbf{w}_t. \tag{5}$$

[4]We use the code made available by Todd Clark.

[5]We assume a Minnesota prior for the VAR dynamic parameters, and the overall tightness prior is chosen by maximizing the marginal data density of the BVAR (with constant variance) as in Carriero et al. (2015). After the overall prior tightness is set, we apply the BVAR-SV estimation algorithm described in Carriero et al. (2022).

Because data revisions may add new information (as reviewed by Mankiw & Shapiro, 1986, and Clements & Galvão, 2019), then $cov(\mathbf{e}_t^{KK}, \mathbf{w}_t)$ is not a zero matrix. Define ζ_t as the Nl vector of disturbances that include the innovations $\mathbf{e}_t^{KK}$ from Equation (3) and $\mathbf{w}_t$ from Equation (5) such that $var(\zeta_t) = \mathbf{Q}$. News revisions then imply that off-diagonal elements of $\mathbf{Q}$ may be nonzero, as the revisions disturbances are correlated with the disturbances to the true values.

KK propose estimating the parameters in Equations (3) and (5) jointly using the seemingly unrelated regression estimator (SUR) with observations up to $T - l + 1$. Then using observations on $\mathbf{z}_t^{t+1}$ for $t = T - l + 2, \dots, T$, filtered values for $\mathbf{y}_t$ for $t = T - l + 2, \dots, T$ are obtained using the Kalman filter. The Kalman filter is applied to a state-space representation of the model with state equations given by (3) and (5), and with the assumption that $\zeta_t \sim N(0, \mathbf{Q})$. The observation equations, assuming that $l > p$ and $q = 1$ to simplify the exposition, are given by

$$\mathbf{z}_t^{t+1} = \mathbf{d} + \begin{bmatrix} \mathbf{I}_{N(l-1)} & \mathbf{I}_{N(l-1)} \\ \mathbf{I}_N & \mathbf{0}_N \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{z}}_t \\ (\mathbf{rev}_t - \mathbf{k}_0) \end{bmatrix}, \tag{6}$$

where the Nl vector of intercepts is $\mathbf{d} = \left[(\mathbf{c}^{KK} + \mathbf{k}_{0,(1)})', \dots, (\mathbf{c}^{KK} + \mathbf{k}_{0,(l-1)})', \mathbf{c}^{KK\prime}\right]'$, where $\mathbf{k}_{0,(l-1)}$ is a $N \times 1$ vector included in $\mathbf{k}_0$ referring to revisions to the $(l-1)^{th}$ estimate, and $\widetilde{\mathbf{z}}_t = \left[\left(\mathbf{y}_t - \mathbf{c}^{KK}\right)', (\mathbf{y}_{t-1} - \mathbf{c}^{KK})', \dots, (\mathbf{y}_{t-l+1} - \mathbf{c}^{KK})'\right]'$.

We propose instead to apply Bayesian methods to estimate the parameters in (3) and (5). In particular, we use Gibbs sampling to obtain posterior densities for $\mathbf{c}^{KK}$, $\boldsymbol{\beta}^{KK}$, $\mathbf{k}_0$, $\mathbf{K}$, and $\mathbf{Q}$ and also for the values of the unobserved state variables $\mathbf{y}_t$ and $\mathbf{rev}_t$ for $t = T - l + 2, \dots, T$. The algorithm to sample draws $j = 1, \dots, J$ for parameters is as follows.

First, we obtain draws for the $\mathbf{c}^{KK,(j)}$, $\boldsymbol{\beta}^{KK,(j)}$, $\mathbf{k}_0^{(j)}$, $\mathbf{K}^{(j)}$, and $\mathbf{Q}^{(j)}$, conditional on the $(j-1)$ draw of the time series of true values $\{\mathbf{y}_t^{(j-1)}\}_{t=1}^T$ and revisions $\{\mathbf{rev}_t^{(j-1)}\}_{t=1}^T$, using the SUR model strategy described by Greenberg ((2013), ch. 10.1). Assume that $s = N + N(l-1)$ is the size of the vector $\xi_t = (\mathbf{y}_t', \mathbf{rev}_t')'$ and that β is $k \times 1$ vector of parameters where $k = (Np+1)N + (N(l-1)q+1)N(l-1)$. Note that β is the vectorized version of parameters in $\mathbf{c}^{KK}$, $\boldsymbol{\beta}^{KK}$, $\mathbf{k}_0$, $\mathbf{K}$. Then we can create a block-diagonal $s \times k$ matrix $\mathbf{X}_t$ such that the system of equations can be written as

$$\xi_t = \mathbf{X}_t \beta + \zeta_t, \; \zeta_t \sim N(0, \mathbf{Q}). \tag{7}$$

Now assume Gaussian $\beta \sim N(\beta_0, B_0)$ and inverse-Wishart priors $Q^{-1} \sim W(v_0, R_0)$.[6] Then the Gibbs draws are obtained using the closed-form conditional densities, that is,

$$\beta^{(j)} \sim N(\beta_1, B_1) \tag{8}$$

where

$$B_1 = \left[\sum\nolimits_{t=p+1}^{T} \mathbf{X}_t' (\mathbf{Q}^{-1})^{(j-1)} \mathbf{X}_t + B_0^{-1}\right]^{-1}$$
$$\beta_1 = B_1 \left[\sum\nolimits_{t=p+1}^{T} \mathbf{X}_t' (\mathbf{Q}^{-1})^{(j-1)} \xi_t + B_0^{-1} \beta_0\right]^{-1},$$

and for the variance-covariance matrix as

$$(\mathbf{Q}^{-1})^{(j)} \sim W(v_1, R_1)$$

where

$$v_1 = v_0 + T$$
$$R_1 = \left[R_0^{-1} + \sum\nolimits_{t=p+1}^{T} (\xi_t - \mathbf{X}_t \beta^{(j)})(\xi_t - \mathbf{X}_t \beta^{(j)})'\right]^{-1}.$$

(For notational convenience, we assume that $p > q$).

[6]The values for β_0 and B_0 are set using a Minnesota prior as in Carriero et al. (2019). We compute the overall prior tightness, λ_1, using the values that maximize the marginal data density of a VAR(p) of the true values as in (3), which are available up to $T - l + 1$. We set $\lambda_2 = 0.9$ and the prior for the intercepts is set to zero with a large variance. The value for v_0 is $s + 2$ and $R_0 = T\Sigma$, where Σ is a diagonal matrix with elements equal to the variance of the residuals of a AR(1) applied to each variable in ξ_t. We use a similar strategy to obtain the priors for the revisions block, as (5) is a VAR(q) with revisions observable up $T - l + 1$.

The next step of the algorithm is to jointly obtain $\{\mathbf{y}_t^{(j)}\}_{t=1}^T$ and $\{\mathbf{rev}_t^{(j)}\}_{t=1}^T$ by using the state-space representation of the model with parameters set as $\mathbf{c}^{KK,(j)}$, $\boldsymbol{\beta}^{KK,(j)}$, $\mathbf{k}_0^{(j)}$, $\mathbf{K}^{(j)}$ and $\mathbf{Q}^{(j)}$. As suggested by Carter and Kohn (1994), we use the Kalman smoother values for the state variables and their variances to draw $\{\mathbf{y}_t^{(j)}\}_{t=1}^T$ and $\{\mathbf{rev}_t^{(j)}\}_{t=1}^T$ from multivariate Gaussian densities.[7]

We run the Gibbs sampler over 10,000 draws, remove the first 50%, and compute multistep forecasts for each one of the kept draws by iteration including draws from the disturbances. We compute point forecasts and forecast uncertainty by iteration as described earlier for the BVAR specifications. The KK BVAR model will deliver multiple-horizon forecasts that vary with the data maturity; that is, the forecast $\mathbf{y}_{T+h|T}^{T+h+1}$ will differ from $\mathbf{y}_{T+h|T}^{T+h+l}$.

2.2.1 | Adding stochastic volatility

A novelty of our paper is to allow for time-varying conditional error variances in the KK model. Our precise empirical implementation is motivated by the properties of the data revisions of the UK and US series we consider (see Section 2.3).

For the KK BVAR model, we add stochastic volatility by assuming that

$$var\begin{bmatrix}\mathbf{e}_t^{KK}\\ \mathbf{w}_t\end{bmatrix} = var(\zeta_t) = \mathbf{Q}_t = \mathbf{A}^{-1}\boldsymbol{\Lambda}_t\mathbf{A}^{-1},$$

where $\boldsymbol{\Lambda}_t$ is a diagonal matrix and $\mathbf{A}^{-1}$ is lower triangular with ones on its main diagonal, as in Carriero et al. (2019). This specification permits time-varying volatilities for the disturbances in the equations for both the true values and the revisions. Allowing for SV in this form implies that

$$\zeta_t = \mathbf{A}^{-1}\boldsymbol{\Lambda}_t^{0.5}\eta_t,\ \eta_t \sim N(0, \mathbf{I}_{Nl}).$$

The fact that $\boldsymbol{\Lambda}_t$ is diagonal implies that the j^{th} element of the rescaled disturbances, $\tilde{\zeta}_t = \mathbf{A}\zeta_t$, can be written as $\tilde{\zeta}_{j,t} = \sqrt{\lambda_{j,t}}\eta_{j,t}$. The observational link between the disturbances of the KK model (Equation 7) and the unobserved volatility processes is

$$\ln\tilde{\zeta}_{j,t}^2 = \ln\lambda_{j,t} + \ln\eta_{j,t}^2. \tag{9}$$

The process for $\lambda_{j,t}$ is given by

$$\ln\lambda_{j,t} = \ln\lambda_{j,t-1} + \epsilon_{j,t}, \tag{10}$$

where $\epsilon_{j,t}$ is the j^{th} element of ϵ_t, and $\epsilon_t \sim N(0, \Phi)$, with no constraints imposed on the variance-covariance matrix Φ.

We estimate the KK BVAR-SV model by adding additional steps to the Gibbs sampler described earlier to implement the Kim et al. (1998) algorithm to draw values for the time-varying volatilities using the state-space form implied by equations (9) and (10). The additional steps draw the values in $\mathbf{A}$ from a a Gaussian density, and Φ from an inverse Wishart, as in Carriero et al. (2019).

A Gibbs sampler is then employed to obtain the posterior distribution of the constant parameters $\mathbf{c}^{KK}$, $\boldsymbol{\beta}^{KK}$, $\mathbf{k}_0$, $\mathbf{K}$, the time-varying variance-covariance matrix $\mathbf{Q}_t$ (which requires draws of $\mathbf{A}$ and Φ), and for the unobserved variables $\mathbf{y}_t$ and $\mathbf{rev}_t$ for $t = T - l + 2, \ldots, T$. The algorithm has then three blocks. In the first block we obtain a draw for the parameters $\mathbf{c}^{KK}$, $\boldsymbol{\beta}^{KK}$, $\mathbf{k}_0$, $\mathbf{K}$ using the SURE strategy described earlier, that is, by obtaining draws for $\beta^{(j)}$ using (8) conditional on $\mathbf{Q}_t^{(j-1)}$, $\{\mathbf{y}_t^{(j-1)}\}_{t=1}^T$, and $\{\mathbf{rev}_t^{(j-1)}\}_{t=1}^T$. In the second block we use the algorithm in Carriero et al. (2019) to obtain draws for $\mathbf{A}$, $\boldsymbol{\Lambda}_t$ and Φ, conditional on $\beta^{(j)}$, $\{\mathbf{y}_t^{(j-1)}\}_{t=1}^T$ and $\{\mathbf{rev}_t^{(j-1)}\}_{t=1}^T$. The third block obtains draws for $\{\mathbf{y}_t^{(j)}\}_{t=1}^T$ and $\{\mathbf{rev}_t^{(j)}\}_{t=1}^T$, conditional on $\beta^{(j)}$ and $\mathbf{Q}_t^{(j)}$, by using a multivariate Gaussian density with mean and variance obtained from the Kalman smoother equations. We run the Gibbs sampler over 10,000 draws, remove the first 50%, and compute multistep forecasts for each one of the kept draws. We compute forecasts by iteration as in Clark and Ravazzolo (2014), so draws from the disturbances are included at each horizon. These draws use $\zeta_t \sim N(0, Q_{T+h|T})$, that is, the variance-covariance matrix may change with the horizon, as we compute $\ln\lambda_{j,T+1}, \ldots, \ln\lambda_{j,T+h}$ using Equation (10) with draws from the variance equation disturbances. We apply a similar approach to calculate forecasts from the BVAR-SV models.

[7] This step makes provision for the fact that the values of these state variables are observed up to $T - l + 1$.

2.3 | Empirical implementation of the KK BVAR approach

The KK was used for forecasting by Kishor and Koenig (2012), Kishor and Koenig (2014), and Clements and Galvão (2013a). But these studies were based on more restricted specifications: They did not allow for stochastic volatility and did consider the case of more than one variable being subject to revisions. In this section, we discuss how we implement the KK BVAR approach, guided by the properties of the data.

The approach is applied to the 4-variable BVAR in Clark and Ravazzolo (2014), who established the benefits of allowing for SV when forecasting with BVAR models. The VAR comprises four key quarterly macroeconomic variables: the first differences of the logs of real GDP and the GDP deflator (so that these variables are effectively growth rates), the unemployment rate and an interest rate. As in Clark and Ravazzolo (2014), we set the autoregressive order of the VAR in (3) to four.

We estimate the KK BVAR model for the United States and the United Kingdom. For the United States, the real-time data for real GDP, the GDP deflator and unemployment are all obtained from the Philadelphia Fed Real-Time Database,[8] and we consider quarterly vintages. Prior to 1991, GDP is in fact GNP. The interest rate is the 3-month Treasury Bill rate. For the United Kingdom, we obtain monthly real-time vintages from the Office of National Statistics (ONS) website on real GDP (GDP in chained volume measures) and on nominal GDP (GDP at current prices).[9] We compute the implied GDP deflator using the ratio between the nominal and real GDP values, and use the monthly vintages that include first releases as the quarterly vintages.[10] UK data on the unemployment rate and the 3-month interbank rate are taken from the St Louis FRED dataset.[11] The sample period starts in 1965Q4 for US data and only in 1990Q1 for UK data, because the UK time series are limited by the availability of the real-time nominal GDP data.

We do not model revisions for all the endogenous variables. The interest rate is not subject to revisions, and revisions to unemployment are small and infrequent. Hence, the vector $\mathbf{y}_t$ is partitioned into N_1 and a N_2 vectors, $\mathbf{y}_t = (y'_{1t}, y'_{2t})'$. The number of equations in the system given by (3) and (5) is then $s = N + (l-1)N_1$, where N_1 is the number of endogenous variables for which we model revisions.

We assume that the true values are observed about 2 years after the first estimate, that is, $y_{1t}^{t+8} = y_{1t}$ as $l = 8$. Figure 1 presents the first two revision processes, that is, $\mathbf{rev}_t^{(1)} = (\mathbf{y}_{1t}^{t+1} - \mathbf{y}_{1t})$ and $\mathbf{rev}_t^{(2)} = (\mathbf{y}_{1t}^{t+2} - \mathbf{y}_{1t})$,[12] for the sample period for which we have y_{1t}^{t+8}. The top panel graphs the revisions to US output growth and inflation, while the bottom panel shows the same for the United Kingdom. Revisions have little first-order serial correlation except for the revisions to UK inflation, which have negative serial correlation of about −0.2.[13] The correlation between US output growth and inflation revisions are large at −28% for revisions to the first estimate, and as much as −48% for revisions to the second estimate. Correlations for the UK are smaller than 5%. The KK BVAR approach allows the innovations to the true values $\mathbf{y}_t$ and the revisions processes $\mathbf{rev}_t$ to be correlated, and the approach is able exploit the empirical correlations between the revisions to output growth and inflation reported here.

Figure 2 presents the standard deviations of these revisions computed over rolling windows of 20 observations, that is, over five year periods, and Figure 3 presents the standard deviations of the 1st, 8th and latest-vintage (2019Q4) estimates. Taken together, these figures suggest a role for stochastic volatility in modelling data revisions. The size of the revisions between the 1st–8th estimates, and the 2nd–8th estimates, fluctuate over the period, and are larger for the US than the UK for growth, whereas for inflation the relative magnitudes are reversed. The variability of revisions to UK inflation display an upward trend, whereas those to US inflation have generally been decreasing. For the US, the variability of the estimates for all maturities has moved closely together over time, on a downward trend, for both variables. There was a sharp increase in the variability of estimates of UK growth in 2008, and a halving in the variability of final-vintage UK estimates around 2008. In addition, these results are consistent with the well-documented Great Moderation in the United

[8] https://www.philadelphiafed.org/research-and-data/real-time-center/real-time-data/.

[9] https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/realtimedatabaseforukgdpabmi and https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/realtimedatabaseforukgdpybha.

[10] We use the implicit deflator here for comparability with US data. The Bank of England targets the CPI measure of inflation, which is not subject to revisions because of UK regulations.

[11] https://fred.stlouisfed.org/.

[12] To ease comparison, we plot $\mathbf{rev}_t^{(2)}$ instead of $\mathbf{rev}_{t-1}^{(2)}$ as it is included in (4).

[13] As indicated in the note to the figure, all the other first-order autocorrelations are no larger in absolute value than 0.12.

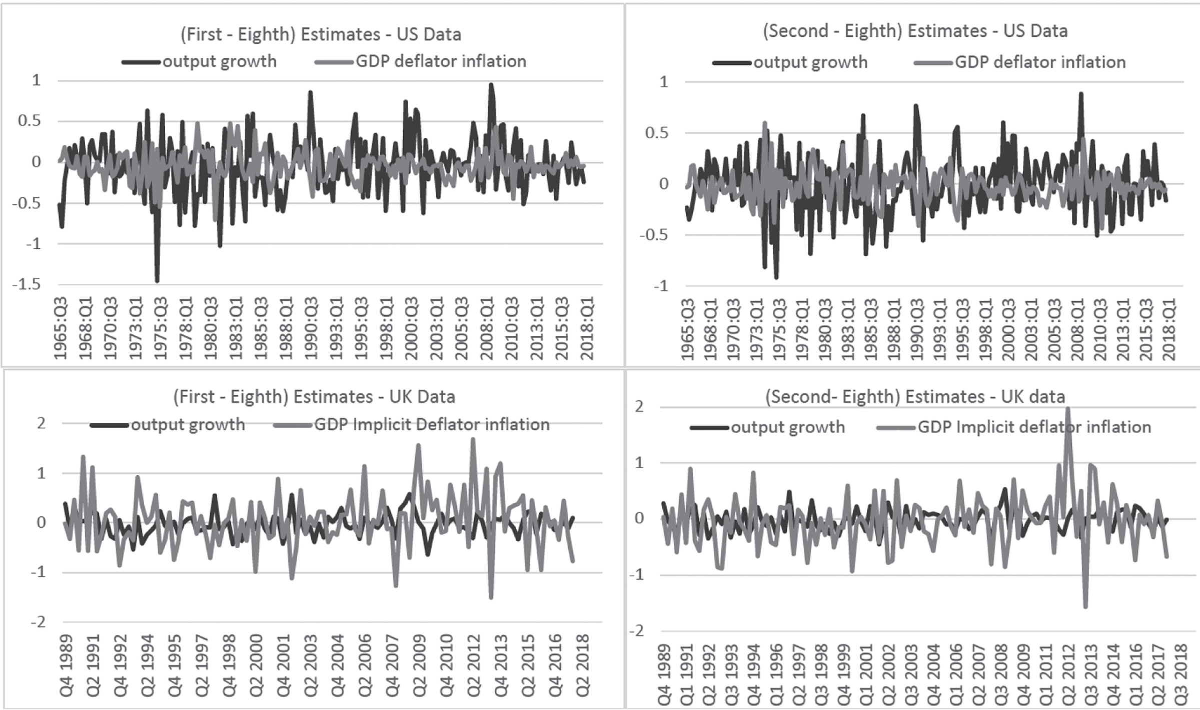

FIGURE 1 Time series of revisions for US and UK first and second estimates assuming the eighth estimates are 'true' values. Notes: Correlation between output and deflator revisions: −28% (1^{st}–8^{th}) and −48% (2^{nd}–8^{th}) for the United States; −2% (1^{st}–8^{th}), and 5% (2^{nd}–8^{th}) for the United Kingdom. First-order autocorrelations for US 1^{st}–8^{th} revisions for output and inflation are 0.01, 0.01, and for 2^{nd}–8^{th} are 0.09 and 0.01. For the United Kingdom, the comparable figures are 0.08, −0.21 (1^{st}–8^{th}) and −0.12, −0.21 (2^{nd}–8^{th})

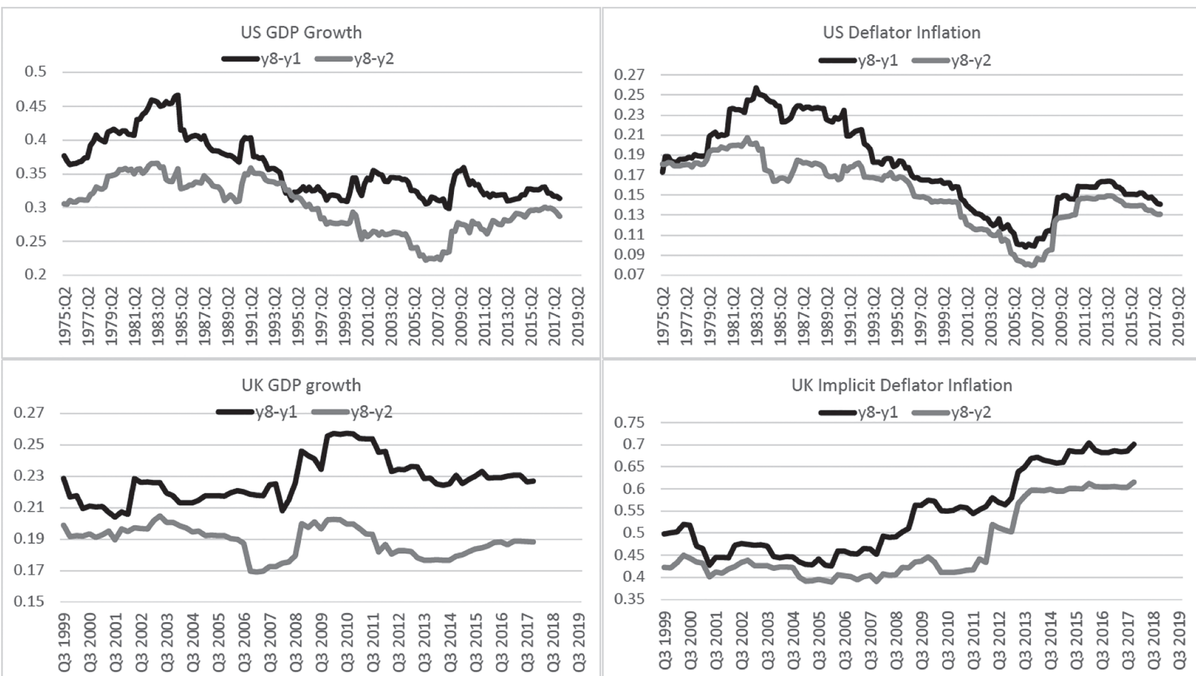

FIGURE 2 Revision size: standard deviation of revisions computed for a rolling window of 5 years. Notes: Date indicates the reference quarter of the last observation in the rolling window

States, as the variance for both output growth and inflation declines in the 80s. The variance of GDP increases in 2008/9 recession. This appears more dramatic for the United Kingdom, but note for the United Kingdom the more volatile period of the previous century is not shown.

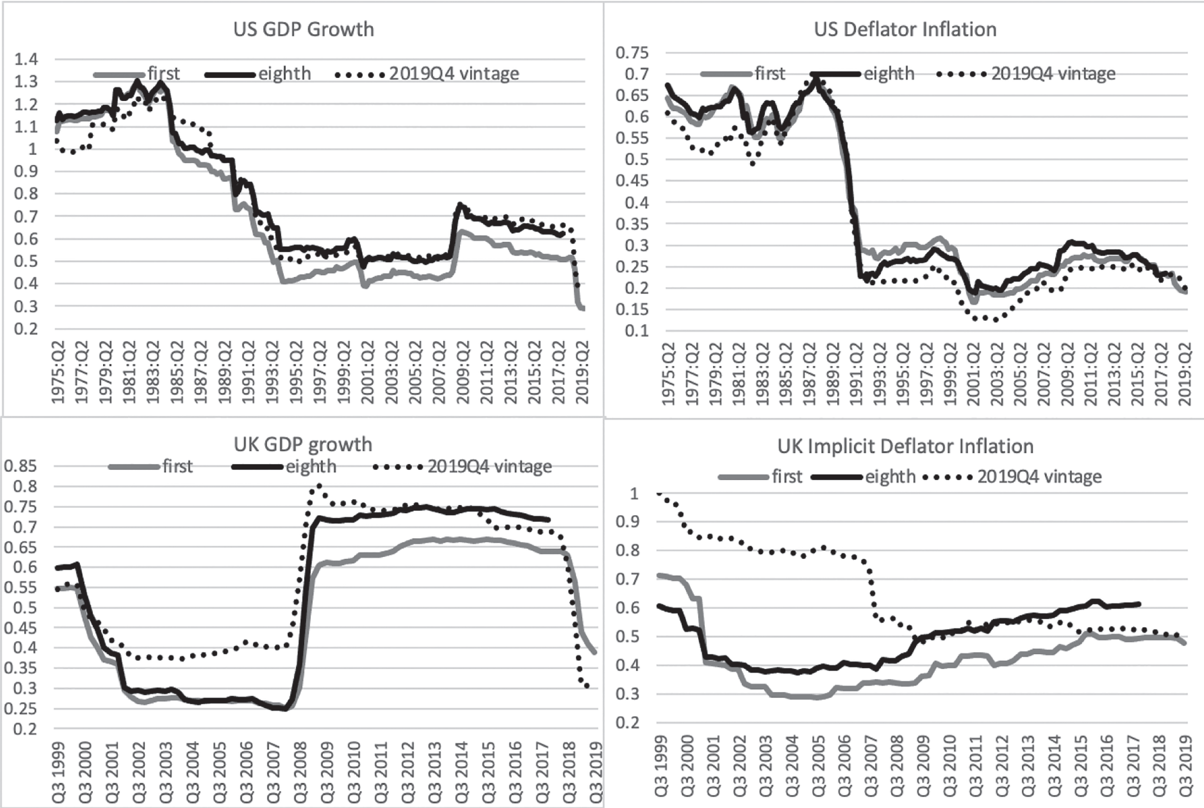

FIGURE 3 Standard deviation of the first, eighth and EOS (2019Q4) estimates computed using rolling windows of 5 years. Notes: Date indicates the reference quarter of the last observation in the rolling window

This initial data analysis indicate that the KK BVAR-SV may be a useful model for real-time forecasting, as the model is able to capture time-variation in the innovations to both true values and revision processes in addition to accommodate the contemporaneous correlation between revisions to different variables. Because our analysis suggests only limited serial correlation in the revisions, we restrict the dynamics in Equation (5) to a single lag.[14] The number of elements of the matrix $\mathbf{K}_1$ in Equation (5) increases quickly with l, as the matrix captures the dynamics of $l-1$ revisions process for each variable in y_{1t} allowing for nonzero cross dynamic correlations. We consider two KK BVAR-SV specifications. The first one assumes $l = 8$, but $\mathbf{K}_1$ is assumed to be diagonal, limiting the scope of the cross variable-revision dynamics as there is limited serial correlation in the data revisions. The second specification assumes that $l = 2$, following Kishor and Koenig (2012) but does not impose any restrictions on $\mathbf{K}_1$. This last specification works well for the US data because initial revisions to output growth and inflation are large relative to subsequent ones, but less well for the UK data where later revisions are more important.[15]

As a consequence, we only apply this last specification to US data. In both specifications, the contemporaneous correlation across revision variables are considered to be nonzero in $\mathbf{Q}_t$.

3 | APPLICATIONS TO US AND UK REAL-TIME DENSITY FORECASTING

In this section, we use BVAR models estimated by both RTV and KK to forecast US and UK macroeconomic variables. Our aims are to assess the relevance of accounting for data uncertainty when making probabilistic forecasts of macrovariables using BVARs and to determine whether one of the two approaches we consider is superior to the other. However, before

[14]Note in the footnote of Figure 1, we have reported serial correlation coefficients for revisions between the 1st–8th estimates, and the 2nd–8th estimates, but would expect serial correlation to matter even less for revisions to later releases (e.g., sth-8th, for $8 > s > 2$) as differences to the true values are frequently zero.

[15]Figure 3 provides evidence supporting this statement. The differences between the characteristics of eighth-release data and the 2019Q4 vintage are large for the UK before 2007.

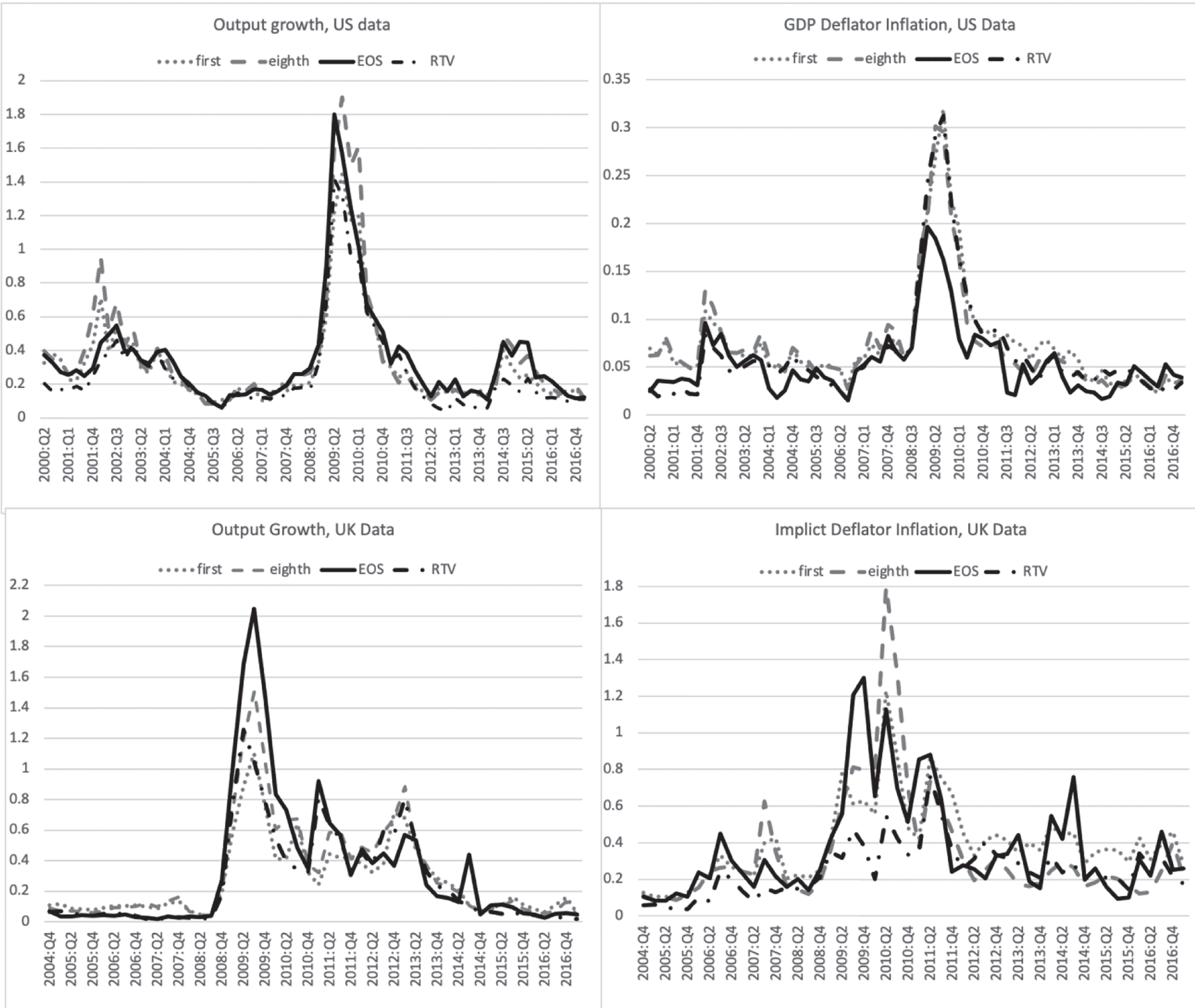

FIGURE 4 One-step-ahead predictive variance (out-of-sample) for three forecasting models: BVAR-SV EOS, BVAR-SV RTV, KK-BVAR-SV ($l = 8$). Notes: The values are the variance of the one-step-ahead predictive density at each forecast origin (computed using 5000 draws from the parameters' posterior at each origin (after the first 5,000 draws were removed as burn-in), where models were re-estimated with increasing windows of data at each origin). The predictive variances of the KK-BVAR-SV model are ‘first’ (first release) and ‘eight’ (last release). RTV and EOS are from a BVAR-SV model estimated with the indicated approach

presenting the forecast comparisons in Section 3.3, Section 3.2 considers the predictive variances. These are a major driver of the results for the densities and provide insights into the effects of the data revisions on density forecast performance.

3.1 | Forecasting exercise design

We set the US out-of-sample period to the forecast origins (vintages) of 2000Q2 to 2017Q1 (a total of 68), while for the United Kingdom, it is 2004Q4 to 2017Q1 (50 origins). These sample periods are constrained by the fact that we look at both one-quarter-ahead and 1-year forecasts that are evaluated using both the first and the eighth estimates. As the latest vintage we have is 2019Q4, with data up to 2019Q3, we can only evaluate forecasts computed up to the 2017Q1 origin. For both countries we use increasing estimation windows (a ‘recursive scheme’), because larger sample sizes are helpful when forecasting with models with SV. Nonetheless, estimation periods are clearly markedly shorter for the United Kingdom.

We evaluate the forecasting accuracy of each BVAR specification for the four variables, for two vintages of actual values. These are: the first release, such that the target is y_{T+1}^{T+2} for one-step-ahead forecasts, and y_{T+4}^{T+5} for four-step-ahead forecasts and the eighth release with 1 and 4-step horizon targets of y_{T+1}^{T+9} and y_{T+4}^{T+12}.

Two measures of forecast performance are reported. For point forecasts, we use the root mean squared forecast error (RMSFE). For density forecast evaluation, we calculate minus the log of the predictive density score (logscore), such that

TABLE 1A Relative density forecasting performance (logscore) of RTV and KK approaches to BVAR benchmark: stochastic volatility specifications, BVAR-SV: US Dataset—2000Q2–2017Q1 (forecasting origins, 68 obs)

	One-quarter-ahead forecasts, $h = 1$				**One-year-ahead forecasts, $h = 4$**			
	EOS	**RTV**	**KK, $l = 2$**	**KK, $l = 8$**	**EOS**	**RTV**	**KK, $l = 2$**	**KK, $l = 8$**
Output growth								
First release	0.781	0.008	−0.073	−0.096	0.919	0.065	−0.041	0.018
		[0.141]	[**−1.338**]	[**−2.185**]		[0.719]	[−0.791]	[0.421]
Eighth release	0.988	0.158	0.004	−0.030	1.119	0.149	0.017	−0.045
		[1.385]	[0.058]	[−0.610]		[1.266]	[0.294]	[**−1.380**]
GDP deflator inflation								
First release	0.141	−0.049	−0.084	−0.112	0.100	0.054	0.006	0.070
		[−0.707]	[−1.077]	[**−1.381**]		[1.470]	[0.158]	[1.871]
Eighth release	0.148	−0.012	−0.111	−0.123	0.130	0.053	0.019	0.056
		[−0.161]	[**−1.422**]	[**−1.403**]		[1.421]	[0.522]	[1.593]
Additional variables								
Unemp	0.127	0.008	−0.060	−0.085	2.307	0.014	−0.352	−0.345
		[0.145]	[−0.672]	[−0.875]		[0.195]	[**−1.572**]	[**−1.686**]
3-mon rate	0.193	0.022	−0.048	0.024	1.805	0.117	−0.053	−0.054
		[0.433]	[−0.589]	[0.577]		[1.162]	[−0.466]	[−1.034]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are the logscore in the EOS (benchmark) column, differences to the logscore in the other columns. Values in brackets and t-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

TABLE 1B Relative density forecasting performance (logscore) of RTV and KK approaches to BVAR benchmark: stochastic volatility specifications, BVAR-SV: UK Dataset—2004Q4–2017Q1 (forecasting origins, 50 obs)

	One-quarter-ahead forecasts, $h = 1$			**One-year-ahead forecasts, $h = 4$**		
	EOS	**RTV**	**KK, $l = 8$**	**EOS**	**RTV**	**KK, $l = 8$**
Output growth						
First release	0.824	−0.012	−0.080	3.548	0.446	−1.568
		[−0.187]	[−0.674]		[1.099]	[**−1.286**]
Eighth release	1.238	0.100	−0.143	4.849	0.629	−1.944
		[0.813]	[−1.193]		[1.156]	[**−1.322**]
Implicit deflator inflation						
First release	0.921	−0.086	−0.068	0.798	0.009	−0.004
		[−0.670]	[−0.610]		[0.073]	[−0.079]
Eighth release	1.097	0.005	0.072	1.044	0.041	−0.024
		[0.030]	[0.885]		[0.408]	[−0.445]
Additional variables						
Unemployment	−0.122	0.139	0.113	1.554	0.068	−0.117
		[1.005]	[1.273]		[0.620]	[−0.592]
3-month rate	0.352	−0.024	−0.114	1.601	0.384	0.092
		[−0.156]	[−0.763]		[2.190]	[0.751]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are the logscore in the EOS (benchmark) column, differences to the logscore in the other columns. Values in brackets and t-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

a smaller (more negative) value is preferred. If $p_{T+h|T}(.)$ is the h-step ahead density of y made at time T, the logscore is $-\ln(p_{T+h|T}(y_{T+h})$ for realization y_{T+h}. We compute the logscore using its closed-form solution for Gaussian densities (as in Clark & Ravazzolo, 2014), using the mean and the variance obtained using the predictive density draws. When the mean and variance are given by $\mu_{T+h|T}$ and $\sigma^2_{T+h|T}$, the (negative) score is

$$-\ln(p_{T+h|T}(y_{T+h})) = \frac{(y_{T+h} - \mu_{T+h|T})^2}{2\sigma^2_{T+h|T}} + \frac{1}{2}\ln(\sigma^2_{T+h|T}) + \frac{1}{2}\ln(2\pi). \tag{11}$$

The negative of the logscore computed analytically for a normal predictive density is equivalent to the Dawid–Sebastiani score function. It is a proper score function, meaning that the optimal forecast is to deliver the true density function—there is no incentive to gameplay (as discussed in Gneiting & Katzfuss, 2014).

We test whether the differences in forecast performance between the models are statistically significant using the Diebold and Mariano (1995) (DM) test statistic. Values of the DM statistic in bold in the tables signify rejection of the null hypothesis of equal accuracy in favour of the alternative (to the benchmark, either the BVAR-SV EOS or BVAR EOS) at the 5% level (with critical values from a Gaussian distribution, and the Newey–West variance estimator).

3.2 | Data uncertainty and the predicted variances

The analytical results in Appendix A suggest that when data revisions are news, we would expect that $\sigma^{2,EOS}_{T+1|T} < \sigma^{2,RTV}_{T+1|T}$. If data revisions are noise, we should observe $\sigma^{2,EOS}_{T+1|T} > \sigma^{2,RTV}_{T+1|T}$. In terms of the predictive densities obtained from the KK BVAR-SV models, we expect that the predictive variance for the first estimate, $\sigma^{2,T+2,KK}_{T+1|T}$, will be lower than the predictive variance for the mature estimate, $\sigma^{2,T+l+1,KK}_{T+1|T}$, if data revisions are news. Opposite inequalities are expected if data revisions are noise. Although the classification of data revisions as news and noise has implications for the predictive variances of the RTV and the KK approaches, these modelling approaches do not require us to explicitly identify the news and noise components, as in Jacobs and van Norden (2011). Nevertheless, in this section, we calculate the time-series of predictive variances from the BVAR-SV over the out-of-sample period, when the model is estimated by EOS, RTV, and KK.

Figure 4 presents the one-step-ahead predictive variances, $\sigma^2_{T+1|T}$ computed at each forecast origin of the out-of-sample period. These are the variances of the one-step-ahead predictive densities obtained by using 5000 draws of the parameters' posteriors (after the first 5000 Gibbs draw are removed) and are calculated using all data up to T (resulting in increasing sample periods at successive forecast origins), by the iterative procedure described in Section 2. We graph the predictive variances for the KK BVAR-SV model ($l = 8$), and the BVAR-SV model estimated with EOS and RTV data.

The variation of the RTV one-step predictive variances for UK inflation is attenuated compared with EOS, in particular during the turbulent 2009–2011 period, when the RTV variances are almost a half of EOS values. This suggests that data revisions are likely to be linked to new information as $\sigma^{2,EOS}_{T+1|T} > \sigma^{2,RTV}_{T+1|T}$. Similar behaviour, but with more modest differences, is observed for US output growth. For UK output growth, we find similar results to US output, except for 2012, when the RTV estimated volatility is higher than EOS. For US inflation, we find that for most time periods, $\sigma^{2,EOS}_{T+1|T} < \sigma^{2,RTV}_{T+1|T}$, compatible with noise revisions. In all cases, sizeable differences in predictive variances are usually observed during turbulent periods (e.g., 2009–2011), suggesting that accounting for data uncertainty may have an episodic impact on the density forecasting performance of the BVAR-SV model.

Turning now to the KK predictive variances which target first or eighth estimates, we find that $\sigma^{2,T+2,KK}_{T+1|T} < \sigma^{2,T+9,KK}_{T+1|T}$ for output growth, which is to be expected if revisions 'add news'. For both inflation series, we find that $\sigma^{2,T+2,KK}_{T+1|T} > \sigma^{2,T+9,KK}_{T+1|T}$ for forecast origins after 2012, which may indicate that data revisions are noise. These results support the view that the KK modelling approach can accommodate both news and noise revisions, that the estimated differences between first and final releases predictive variance is usually moderate, but that the sign of their differences may change over time.

3.3 | Relative forecasting performance

As the literature suggests the inclusion of SV improves density forecasting (see Clark, 2011; Clark & Ravazzolo, 2014; Diebold et al., 2017), our first set of results evaluate the impact of data uncertainty by comparing the performance of the RTV and KK approaches against EOS for BVAR-SV models. We then contrast these results to those for constant volatility specifications. Tables 1A and 1B show results using the logscore to measure density forecasting accuracy, with Tables 2A and 2B presenting equivalent results for point forecasts using RMSFE as the loss function. The tables record results for both datasets (United States and United Kingdom), two horizons ($h = 1, 4$), and two data maturities (first and eighth). For the US data, we consider two KK specifications ($l = 2$ and $l = 8$). The tables include the t-statistics of the DM test of equal accuracy against the benchmark (EOS).

We find no evidence of statistically significant improvements in density forecasting performance from estimating BVAR-SV models by RTV, relative to EOS. However, the KK ($l = 8$) approach does improve the accuracy of BVAR-SV

TABLE 2A Relative point forecasting performance (RMSFE) of RTV and KK approaches to BVAR benchmark: stochastic volatility specifications, BVAR-SV: US Dataset—2000Q2–2017Q1 (forecasting origins, 68 obs)

	One-quarter-ahead forecasts, $h = 1$				One-year-ahead forecasts, $h = 4$			
	EOS	**RTV**	**KK, $l = 2$**	**KK, $l = 8$**	**EOS**	**RTV**	**KK, $l = 2$**	**KK, $l = 8$**
Output growth								
First release	0.558	0.968	0.951	0.937	0.622	0.988	0.957	0.956
		[−0.887]	[−0.881]	[**−1.741**]		[−0.524]	[−0.970]	[−1.217]
Eighth release	0.626	0.987	0.993	0.972	0.687	1.000	0.995	0.976
		[−0.489]	[−0.200]	[−1.126]		[0.004]	[−0.118]	[−0.757]
GDP deflator inflation								
First release	0.250	1.007	0.989	1.024	0.265	1.004	0.996	1.022
		[0.237]	[−0.282]	[0.732]		[0.116]	[−0.176]	[0.329]
Eighth release	0.258	1.022	0.990	1.030	0.279	1.027	1.006	1.043
		[0.966]	[−0.233]	[0.809]		[0.811]	[0.344]	[1.419]
Additional variables								
Unemp	0.283	1.031	1.025	1.100	1.078	1.006	0.971	1.017
		[1.248]	[0.883]	[1.484]		[0.460]	[−1.272]	[0.925]
3-month rate	0.385	0.977	0.971	0.946	1.352	1.036	1.003	0.985
		[−0.462]	[−0.775]	[−1.170]		[1.104]	[0.114]	[−0.456]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are RMSFE in the EOS (benchmark) column, ratios to the benchmark RMSFE in the other columns. Values in brackets and t-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

TABLE 2B Relative point forecasting performance (RMSFE) of RTV and KK approaches to BVAR benchmark: stochastic volatility specifications, BVAR-SV: UK Dataset—2004Q4–2017Q1 (forecasting origins, 50 obs)

	One-quarter-ahead forecasts, $h = 1$			One-year-ahead forecasts, $h = 4$		
	EOS	**RTV**	**KK, $l = 8$**	**EOS**	**RTV**	**KK, $l = 8$**
Output growth						
First release	0.613	0.952	0.864	0.747	1.083	0.875
		[−0.760]	[**−1.359**]		[1.355]	[**−2.142**]
Eighth release	0.685	0.920	0.871	0.802	1.047	0.909
		[−1.296]	[**−1.538**]		[1.113]	[**−2.346**]
Implicit deflator inflation						
First release	0.626	0.833	0.942	0.536	0.997	0.946
		[**−1.957**]	[−0.638]		[−0.099]	[−0.797]
Eighth release	0.683	0.871	1.021	0.582	1.021	1.028
		[**−1.382**]	[0.383]		[0.789]	[0.557]
Additional variables						
Unemployment	0.232	1.039	1.055	0.794	1.047	1.046
		[0.838]	[1.079]		[1.563]	[1.433]
3-month rate	0.425	0.999	0.930	1.310	1.224	0.999
		[−0.006]	[**−1.312**]		[2.386]	[−0.019]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are RMSFE in the EOS (benchmark) column, ratios to the benchmark RMSFE in the other columns. Values in brackets and t-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

density forecasts, in particular the one-step-ahead forecasts of output growth and inflation for the US data, and the $h = 4$ UK output growth forecasts. There is less evidence that KK is beneficial for point forecasting, for the United States, although for the UK KK generates more accurate $h = 4$ forecasts, and RTV more accurate one-step-ahead points forecasts.

Although our main focus is on SV models, the results for constant volatility models are illuminating. Tables 3A and 3B report results for point forecasting, and Tables 4A and 4B result for density forecasting, both for constant-volatility

TABLE 3A Relative point forecasting performance (RMSFE) of RTV and KK approaches to BVAR benchmark: constant volatility specifications, BVAR: US Dataset—2000Q2–2017Q1 (forecasting origins, 68 obs)

	One-quarter-ahead forecasts, $h = 1$				One-year-ahead forecasts, $h = 4$			
	EOS	RTV	KK, $l = 2$	KK, $l = 8$	EOS	RTV	KK, $l = 2$	KK, $l = 8$
Output growth								
First release	0.644	0.924	0.973	0.875	0.729	0.943	0.985	0.908
		[**−1.575**]	[−0.380]	[**−2.752**]		[**−2.531**]	[−0.703]	[**−1.993**]
Eighth release	0.691	0.965	1.039	0.957	0.787	0.950	1.012	0.954
		[−1.059]	[0.741]	[**−1.413**]		[**−2.261**]	[0.480]	[−1.122]
GDP deflator inflation								
First release	0.277	0.952	0.937	0.913	0.304	0.907	0.913	0.879
		[**−2.043**]	[**−1.705**]	[**−2.471**]		[**−2.372**]	[**−1.727**]	[**−2.159**]
Eighth release	0.281	0.942	0.911	0.894	0.315	0.897	0.901	0.878
		[**−2.006**]	[**−2.100**]	[**−2.591**]		[**−2.269**]	[**−1.894**]	[**−2.101**]
Additional variables								
Unemp	0.293	1.069	1.055	1.031	1.174	1.042	1.050	1.002
		[1.460]	[1.300]	[1.124]		[2.529]	[2.315]	[0.086]
3-month rate	0.458	0.908	0.912	0.889	1.365	1.002	1.004	1.003
		[−1.268]	[−0.954]	[**−1.360**]		0.067	[0.095]	[0.087]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are RMSFE in the EOS (benchmark) column, ratios to the benchmark RMSFE in the other columns. Values in brackets and *t*-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

TABLE 3B Relative point forecasting performance (RMSFE) of RTV and KK approaches to BVAR benchmark: constant volatility specifications, BVAR: UK Dataset—2004Q4–2017Q1 (forecasting origins, 50 obs)

	One-quarter-ahead forecasts, $h = 1$			One-year-ahead forecasts, $h = 4$		
	EOS	RTV	KK, $l = 8$	EOS	RTV	KK, $l = 8$
Output growth						
First release	0.571	0.933	0.957	0.705	0.989	0.970
		[**−1.561**]	[−0.817]		[−0.285]	[−0.384]
Eighth release	0.655	0.906	0.947	0.765	0.979	1.006
		[**−2.002**]	[−1.250]		[−0.572]	[0.096]
Implicit deflator inflation						
First release	0.598	0.955	0.993	0.607	0.996	0.925
		[−0.764]	[−0.105]		[−0.110]	[−1.303]
Eighth release	0.592	1.024	1.100	0.610	1.035	1.025
		[0.447]	[1.076]		[1.308]	[0.527]
Additional variables						
Unemployment	0.231	1.027	1.072	0.801	0.999	1.089
		[0.976]	[2.063]		[−0.042]	[1.851]
3-month rate	0.504	0.881	0.901	1.516	0.951	1.017
		[−1.052]	[−0.962]		[−0.959]	[0.137]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are RMSFE in the EOS (benchmark) column, ratios to the benchmark RMSFE in the other columns. Values in brackets and *t*-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

BVARs. In line with the empirical results in Clements and Galvão (2013b) and Clements (2017) for autoregressive models, we find that RTV and KK improve the forecasting performance of the BVAR point forecasts for output growth and inflation for the United States (Table 3A). There are also some improvements to US and UK output growth density forecasts (Tables 4A and 4B) for KK, but also now for RTV. Hence contrasting the findings for the specifications with, and without, SV, a clear difference is that RTV loses its value for density forecasting when the models exhibit SV. In the next section, we investigate this using Monte Carlo, to better understand the factors that contribute to modelling data revisions adding value.

TABLE 4A Relative density forecasting performance (logscore) of RTV and KK approaches to BVAR benchmark: constant volatility specifications: US Dataset—2000Q2–2017Q1 (forecasting origins, 68 obs)

	One-quarter-ahead forecasts, $h = 1$				One-year-ahead forecasts, $h = 4$			
	EOS	RTV	KK, $l = 2$	KK, $l = 8$	EOS	RTV	KK, $l = 2$	KK, $l = 8$
Output growth								
First release	0.997	−0.098	−0.035	−0.122	1.121	−0.063	−0.001	−0.067
		[−**1.575**]	−0.853	[−**3.355**]		[−**1.741**]	[−0.065]	[−**2.126**]
Eighth release	1.062	−0.040	0.024	−0.021	1.194	−0.048	0.014	−0.020
		[−0.598]	[0.585]	[−0.886]		[−1.010]	[0.855]	[−0.738]
GDP deflator inflation								
First release	0.119	−0.016	0.004	0.004	0.276	−0.022	0.042	0.036
		[−0.464]	[0.012]	[0.075]		[−0.938]	[1.328]	[1.077]
Eighth release	0.137	−0.029	−0.031	−0.032	0.300	−0.034	0.025	0.023
		[−0.626]	[−0.484]	[−0.490]		[−1.105]	[0.693]	[0.623]
Additional variables								
Unemp	0.203	0.050	0.013	0.004	2.134	0.120	−0.049	−0.252
		[1.031]	[0.205]	[0.056]		[2.652]	[−0.608]	[−**1.290**]
3-month rate	0.791	−0.061	−0.066	−0.021	1.765	−0.002	0.006	0.034
		[−**2.529**]	[−**2.186**]	[−0.842]		[−0.097]	[0.217]	[1.647]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are the logscore in the EOS (benchmark) column, differences to the logscore in the other columns. Values in brackets and *t*-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

TABLE 4B Relative density forecasting performance (logscore) of RTV and KK approaches to BVAR benchmark: constant volatility specifications: UK Dataset—2004Q4–2017Q1 (forecasting origins, 50 obs)

	One-quarter-ahead forecasts, $h = 1$			One-year-ahead forecasts, $h = 4$		
	EOS	RTV	KK, $l = 8$	EOS	RTV	KK, $l = 8$
Output growth						
First release	0.948	−0.066	−0.126	1.917	−0.231	−0.453
		[−0.792]	[−1.031]		[−**1.300**]	[−**1.460**]
Eighth release	1.328	−0.124	−0.195	2.412	−0.302	−0.479
		[−1.123]	[−**1.405**]		[−**1.389**]	[−**1.323**]
Implicit deflator inflation						
First release	0.888	−0.070	−0.023	0.871	0.006	−0.055
		[−1.197]	[−0.406]		[0.103]	[−0.818]
Eighth release	0.886	0.026	0.069	0.950	0.025	−0.023
		[0.389]	[0.840]		[0.609]	[−0.428]
Additional variables						
Unemployment	−0.047	0.072	0.047	1.553	0.111	−0.125
		[1.500]	[1.125]		[1.549]	[−0.878]
3-month rate	0.693	−0.093	−0.009	1.838	−0.003	−0.038
		[−1.300]	[−0.115]		[−0.058]	[−0.568]

Note: Models are re-estimated at each new quarterly forecasting origin by extending the sample period that starts in 1965Q3 for the United States and 1989Q4 for the United Kingdom. Entries are the logscore in the EOS (benchmark) column, differences to the logscore in the other columns. Values in brackets and *t*-statistics of the DM test with the null that the alternative model is as accurate as the benchmark. Values in bold suggest rejection at 10% in a one-sided test with the indicated model under the alternative (CV = −1.282).

4 | A MONTE CARLO EXERCISE

In this section, we report the results of a Monte Carlo exercise designed to further explore when we might expect gains from allowing for data uncertainty when forecasting with BVAR models in real time. The exercise focuses on RTV versus EOS, given that RTV is a simple alternative to EOS, whereas KK is more complicated and entails a number of modelling choices, as illustrated in Section 2. We consider the effects of the forecast horizon and forecast target on the relative performance of the two approaches. As in the empirical exercise, we measure forecast performance using the RMSFE and the logscore, and we evaluate relative forecasting performance using the Diebold–Mariano test for equal forecasting accuracy with the BVAR EOS. We also consider the BVAR-SV as forecasting model to account for the fact that SV may accommodate the effect of different data maturities in the EOS data as described below.

The data generation process (DGP) is the KK model estimated on US data for output growth and GDP deflator inflation. We suppose y_t^{t+14} reveals the true values, so that the DGP includes the initial revisions but also the subsequent annual rounds of revisions.[16] An advantage of having the KK model as the DGP is that data revisions may be a combination of news and noise revisions. We consider a VAR(1) for the $(l-1) = 13$ data revisions processes for each variable (output growth and inflation), and we allow revisions across variables and maturity to be contemporaneously and dynamically correlated. For the true values, we consider a VAR(4), and we estimate the KK VAR model with data from 1985, that is, for the period after the Great Moderation. This lends support to the assumption that the volatility of the disturbances is constant, that is, we do not allow for SV in the DGP.

Using data generated from the KK model, the end-of-sample vintage available at the forecasting origin $T+1$ $\{\mathbf{y}_t^{T+1}\}_{t=1}^{t=T}$ is given by $\mathbf{y}_t^{T+1} = \mathbf{y}_t$ for $t = 1, \ldots, T-l+1$, then $\mathbf{y}_{T-l+2}^{T+1} = \mathbf{y}_{T-l+2} + \mathbf{rev}_{T-l+2}^{(l-1)}$ and so on up to $\mathbf{y}_T^{T+1} = \mathbf{y}_T + \mathbf{rev}_T^{(1)}$. This implies that the variability of the most recent $l-1$ observations which are still subject to revision will differ from that of the fully-revised earlier values. As a consequence, it may be that allowing for stochastic volatility in the BVAR estimated on $\{\mathbf{y}_t^{T+1}\}_{t=1}^{t=T}$ will result in better forecasts, even though stochastic volatility is not intrinsic to the DGP. The rationale is that SV may help to accommodate the changing data variability due to the different maturities of the observations. In contrast to the empirical exercise, the simulation allows us to remove the potentially confounding effects of time-varying volatilities in the process for the true values and data revisions, to isolate the effects of data maturities on relative forecast performance.

For comparison purposes, we consider both the BVAR-SV EOS and BVAR-SV RTV, in addition to the BVAR RTV. We consider a relatively short sample period, matching that of the empirical exercise. The in-sample period $T = 150$, and the out-of-sample period is set to $P = 50$.[17] We evaluate forecasts for both first (y_{T+h}^{T+h+1}) and final releases (y_{T+h}), and for two horizons: one-step-ahead and four-steps-ahead. We set the number of draws to approximate the posterior distribution and the predictive density to 8000 (as first 4000 are removed as burn-in) and the number of Monte Carlo replications is 288. The number of replications is small because of the computational time required to re-estimate each forecasting model with increasing samples over the out-of-sample period, P. Our estimates for accuracy measures (RMSFE and logscore) are actually over 50×288 realizations. The number of replications is used to calculate the rejection rates for the DM statistic. On each replication, DM is computed using $P-h$ observations, as a 5% level one-sided test against the BVAR EOS. The percentage of rejections across replications estimates the rejection frequency against equality with the BVAR EOS, the same benchmark as in the empirical exercises. We also use the results across replications to count the number of times that one of the candidate models is better than the benchmark using either the RMSE or the logscore.

Tables 5A and 5B present the results of the Monte Carlo exercise. As in the case of the empirical exercises, the values for the BVAR EOS column are either the RMSFE or the logscore. For the remaining models, we report ratios to these values (RMSFE) or differences with (logscore).

The results in Tables 5A and 5B suggest that we are more likely to find significant improvements from RTV over EOS for one-step-ahead forecasts, evaluated against the first estimate of output growth. Gains from RTV decline with the horizon. The gains for inflation are in general less marked than for output growth. The allowance for stochastic volatility tends to worsen forecasting performance of one-step-ahead forecasts. But for four-step-ahead forecasts, we find a large proportion of replications (more than 50%) where the inclusion stochastic volatility (as in the BVAR-SV EOS) is beneficial, when forecasts are evaluated against first releases. Indeed at $h = 4$, the BVAR-SV EOS has a similar performance to the BVAR (-SV) RTV.

These results lend some support to the claim that models with stochastic volatility may capture some of the data maturity effects for forecasting the first estimate at longer horizons. They also indicate that RTV is more likely to be beneficial for one-step-ahead point forecasting of initial estimates.

[16]See, for example, Fixler et al. (2014) on US data releases.

[17]At each replication, we simulate $150 + 50 + 100 = 300$ observations and discard the first 100, so as to remove the effects of the initial values on our results.

TABLE 5A Simulation exercise for $T = 150$ and $P = 50$ with KK BVAR(4) $l = 14$ as DGP: One-step-ahead forecasts

	Forecasting first estimate				**Forecasting final estimate**			
	BVAR EOS	BVAR RTV	BVAR-SV EOS	BVAR-SV RTV	BVAR EOS	BVAR RTV	BVAR-SV EOS	BVAR-SV RTV
	RMSFE	RMSFE ratios [% RMSFE ↓] (DM-stat 5% rejections)			RMSFE	RMSFE ratios [% RMSFE ↓] (DM-stat 5% rejections)		
Growth	0.700	0.896 [96%] (**0.472**)	1.140 [22%] (0.008)	0.882 [93%] (0.412)	0.769	0.939 [93%] (**0.588**)	1.081 [1%] (0.000)	0.929 [74%] (0.164)
inflation	0.336	0.984 [62%] (**0.124**)	1.030 [27%] (0.012)	0.974 [60%] (0.116)	0.310	0.988 [73%] (**0.164**)	1.030 [24%] (0.000)	0.975 [48%] (0.032)
	Logsc	Logscore differences [% logscore ↓] (DM-stat 5% rejections)			Logsc	Logscore differences [% logscore ↓] (DM-stat 5% rejections)		
Growth	1.069	−0.108 [86%] (**0.252**)	0.219 [27%] (0.016)	−0.163 [83%] (0.212)	1.166	−0.035 [68%] (**0.204**)	0.110 [6%] (0.000)	−0.023 [34%] (0.036)
inflation	0.371	−0.051 [66%] (**0.160**)	0.088 [24%] (0.012)	−0.037 [60%] (0.104)	0.269	−0.017 [60%] (**0.084**)	0.100 [12%] (0.000)	−0.055 [38%] (0.016)

Note: The entries for BVAR EOS are the average (across replications and out-of-sample period P) RMSFE or logscore. The entries for the other models are RMSFE ratios or differences in logscore. DM t-statistics are computed for each replication using $P - h$ recursive forecasts (computing by re-estimating the forecasting models with increasing samples from $T + 1$ up to $T + P - h$), and entries in () are rejection rates of a 5% sized (one-sided) test in favour of the alternative model. Values in [] are the proportion of the replications where the alternative model has improved accuracy (either RMSFE or logscore) in comparison with the benchmark (BVAR EOS). The DGP is a KK-VAR(4) for output growth and inflation estimated using US data from 1985 and assuming $l = 14$. We compute forecasts using 5000 draws of the posterior distribution for each model after the initial 5000 Gibbs draws were removed as burn-in. Number of Monte Carlo replications is 288.

TABLE 5B Simulation exercise for $T = 150$ and $P = 50$ with KK BVAR(4) $l = 14$ as DGP: Four-step-ahead forecasts

	Forecasting first estimate				**Forecasting final estimate**			
	BVAR EOS	BVAR RTV	BVAR-SV EOS	BVAR-SV RTV	BVAR EOS	BVAR RTV	BVAR-SV EOS	BVAR-SV RTV
	RMSFE	RMSFE Ratios [% RMSFE ↓] (DM-stat 5% rejections)			RMSFE	RMSFE Ratios [% RMSFE ↓] (DM-stat 5% rejections)		
growth	0.680	0.989 [63%] (0.056)	1.006 [61%] (**0.108**)	0.994 [64%] (0.084)	0.782	0.981 [76%] (**0.276**)	1.014 [1%] (0.000)	0.985 [48%] (0.052)
inflation	0.356	1.012 [52%] (0.120)	0.988 [53%] (0.064)	1.001 [58%] (**0.160**)	0.331	1.036 [50%] (0.052)	0.964 [17%] (0.000)	1.025 [36%] (0.024)
	Logsc	Logscore Differences [% logscore ↓] (DM-stat 5% rejections)			Logsc	Logscore Differences [% logscore ↓] (DM-stat 5% rejections)		
growth	1.064	−0.029 [67%] (0.108)	0.180 [60%] (0.112)	−0.137 [69%] (0.108)	1.184	0.008 [52%] (**0.112**)	0.090 [7%] (0.000)	−0.032 [28%] (0.044)
inflation	0.409	0.010 [45%] (0.096)	0.066 [50%] (0.088)	−0.045 [47%] (0.096)	0.332	0.049 [29%] (0.028)	0.066 [10%] (0.000)	−0.047 [24%] (0.008)

Note: The entries for BVAR EOS are the average (across replications and out-of-sample period P) RMSFE or logscore. The entries for the other models are RMSFE ratios or differences in logscore. DM t-statistics are computed for each replication using P-h recursive forecasts (computing by re-estimating the forecasting models with increasing samples from $T+1$ up to $T+P-h$), and entries in () are rejection rates of a 5% sized (one-sided) test in favour of the alternative model. Values in [] are the proportion of the replications where the alternative model has improved accuracy (either RMSFE or logscore) in comparison with the benchmark (BVAR EOS). The DGP is a KK-VAR(4) for output growth and inflation estimated using US data from 1985 and assuming $l = 14$. We compute forecasts using 5000 draws of the posterior distribution for each model after the initial 5000 Gibbs draws were removed as burn-in. Number of Monte Carlo replications is 288.

5 | CONCLUSIONS

In this paper, we consider whether it is possible to improve on the standard practice of effectively ignoring data uncertainty when generating density forecasts from Bayesian VAR models. By 'data uncertainty', we mean that the recent observations at the time a forecast is made will be subject to future data revisions. Such observations are therefore uncertain. Two methods are considered as offering potential improvements—the use of real-time-vintage data (RTV), and simultaneously modelling data revision along with the true or fully revised values of the data. The first is simple to implement. The second requires a Bayesian implementation of the approach of Kishor and Koenig (2012) to allow for stochastic volatility, as well as to allow for the increased number of parameters when we model the multiple rounds of revisions that arise with quarterly data. We provide such an approach, and detail its empirical implementation guided by the properties of the data revisions to the series.

We explore the forecast performances of these approaches, relative to the conventional approach, applied to small VAR models with stochastic volatility, for output growth and inflation for the United States and the United Kingdom. We find that accounting for data uncertainty via the second method—the Bayesian implementation of the approach of Kishor and Koenig (2012), improves the density forecasting performance of BVAR-SV models in some instances: chiefly for the United States. Perhaps just as usefully, we show when allowing for data uncertainty might not work. We find that the extensive point forecasting gains from using RTV in constant-volatility models does not carry over to BVAR-SV models.

A Monte Carlo allows us to investigate the factors which might affect the relative performance of RTV and EOS, and the role played by stochastic volatility when data uncertainty is ignored. We find that RTV frequently improves one-step-ahead BVAR forecasts, but is less successful for four-step-ahead forecasts. Allowing for stochastic volatility in the BVAR improves one-year-ahead forecasts of first releases, even in the absence of underlying stochastic volatility. The BVAR-SV captures the heteroscedasticity resulting from the different data maturities at the end of the estimation sample when the traditional approach to real-time forecasting is used.

ACKNOWLEDGMENTS

Helpful comments from three referees of this journal are gratefully acknowledged.

OPEN RESEARCH BADGES

This article has been awarded Open Data Badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. Data is available at http://qed.econ.queensu.ca/jae/datasets/clements010/.

DATA AVAILABILITY STATEMENT

The data necessary to reproduce the reported results is available in the journal data archive.

ORCID

Michael P. Clements https://orcid.org/0000-0001-6329-1341
Ana Beatriz Galvão https://orcid.org/0000-0003-3263-9450

REFERENCES

Carriero, A., Chan, J., Clark, T. E., & Marcellino, M. (2022). Corrigendum to "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors" [J. Econometrics 212 (1) (2019) 137 154]. *Journal of Econometrics*, *227*(2), 506–512.

Carriero, A., Clark, T. E., & Marcellino, M. (2015). Bayesian VARs: Specification choices and forecast accuracy. *Journal of Applied Econometrics*, *30*(1), 46–73.

Carriero, A., Clark, T. E., & Marcellino, M. (2019). Large vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, *212*, 137–154.

Carriero, A., Clark, T. E., & Marcellino, M. (2020). Capturing macroeconomic tail risks with Bayesian vector autoregressions. Federal Reserve Bank of Cleveland, Working Paper No. 20-02.

Carriero, A., Galvão, A. B., & Kapetanios, G. (2019). A comprehensive evaluation of macroeconomic forecasting methods. *International Journal of Forecasting*, *35*(4), 1226–1239.

Carter, C. K., & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, *81*(3), 541–553.

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

Clark, T. E., & Ravazzolo, F. (2014). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, *30*(4), 551–575.
Clements, M. P. (2017). Assessing macro uncertainty in real-time when data are subject to revision. *Journal of Business & Economic Statistics*, *3*, 420–433.
Clements, M. P., & Galvão, A. B. (2013a). Forecasting with vector autoregressive models of data vintages: US output growth and inflation. *International Journal of Forecasting*, *29*(4), 698–714. https://doi.org/10.1016/j.ijforecast.2011.09.003
Clements, M. P., & Galvão, A. B. (2013b). Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. *Journal of Applied Econometrics*, *28*(3), 458–477. https://doi.org/10.1002/jae.2274
Clements, M. P., & Galvão, A. B. (2019). Data revisions and real-time forecasting. The Oxford Research Encyclopedia of Economics and Finance, Oxford University Press, https://doi.org/10.1093/acrefore/9780190625979.013.248
Corradi, V., & Swanson, N. R. (2006). Predictive density evaluation. In Elliott, G., Granger, C. W. J., & Timmermann, A. (Eds.), *Handbook of economic forecasting, volume 1. Handbook of economics 24*. Horth-Holland: Elsevier, pp. 197–284.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*, 253–263.
Diebold, F. X., Schorfheide, F., & Shin, M. (2017). Real-time forecast evaluation of DSGE models with stochastic volatility. *Journal of Econometrics*, *201*(2), 322–332.
Doan, T., Litterman, R., & Sims, C. A. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, *3*, 1–100.
Fixler, D. J., Greenaway-McGrevy, R., & Grimm, B. T. (2014). The revisions to GDP, GDI, and their major components. *Survey of Current Business*, *August*, 1–23.
Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, *1*, 125–51.
Greenberg, E. (2013). *Introduction to Bayesian econometrics* (2nd ed.). Cambridge University Press.
Jacobs, J. P. A. M., & van Norden, S. (2011). Modeling data revisions: Measurement error and dynamics of 'true' values. *Journal of Econometrics*, *161*, 101–109.
Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies*, *65*(3), 361–393.
Kishor, N. K., & Koenig, E. F. (2012). VAR estimation and forecasting when data are subject to revision. *Journal of Business and Economic Statistics*, *30*(2), 181–190.
Kishor, N. K., & Koenig, E. F. (2014). Credit indicators as predictors of economic activity: A real-time VAR analysis. *Journal of Money, Credit and Banking*, *46*(2-3), 545–564.
Koenig, E. F., Dolmas, S., & Piger, J. (2003). The use and abuse of real-time data in economic forecasting. *The Review of Economics and Statistics*, *85*(3), 618–628.
Mankiw, N. G., & Shapiro, M. D. (1986). News or noise: An analysis of GNP revisions. Survey of Current Business (May 1986), US Department of Commerce, Bureau of Economic Analysis, 20–25.
Rossi, B., & Sekhposyan, T. (2013). Conditional predictive density evaluation in the presence of instabilities. *Journal of Econometrics*, *177*(2), 199–212.
Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*, 1–48.

How to cite this article: Clements M. P., & Galvão A. B. (2022). Density forecasting with Bayesian Vector Autoregressive models under macroeconomic data uncertainty. *Journal of Applied Econometrics*, 1–22. https://doi.org/10.1002/jae.2944

APPENDIX A: TECHNICAL APPENDIX

In the data generating process, the true values y_t follow an AR(1):

$$y_t = \phi y_{t-1} + \eta_t + \nu_t,\ |\phi| < 1, \tag{12}$$

where η_t is the underlying disturbance, and ν_t is a news revision, and the first estimate is given by

$$y_t^{t+1} = y_t - \nu_t + \varepsilon_{t,} \tag{13}$$

with $y_t^{t+n} = y_t$ for $n = 2, 3, \ldots$. Further,

$$\begin{aligned} &\eta_t = \sigma_\eta \xi_{1t};\ \nu_t = \sigma_\nu \xi_{2t};\ \varepsilon_t = \sigma_\varepsilon \xi_{3t}, \\ &\xi_{it} \sim iidN(0,1) \text{ for } i = 1,2,3. \end{aligned} \tag{14}$$

Let $y_t^{t+1} = y_t + rev_t$, and define δ as the relative size of the data revision process, that is,

$$\delta = \frac{var(rev_t)}{\sigma_\eta^2},$$

implying that if data revisions are news, $\delta = \sigma_v^2/\sigma_\eta^2$, and if data revisions are noise, $\delta = \sigma_\varepsilon^2/\sigma_\eta^2$.

A.1 | News revisions

The EOS and RTV forecasts of the mean and the variance are given by

$$\begin{aligned} \mu_{T+1|T}^{EOS} &= \phi y_T^{T+1} = \phi(y_T - \sigma_v \xi_{2T}), \\ \sigma_{T+1|T}^{2,EOS} &= \sigma_\eta^2 + \sigma_v^2 = \sigma_\eta^2(1+\delta), \\ \mu_{T+1|T}^{RTV} &= \phi y_T^{T+1} = \phi(y_T - \sigma_v \xi_{2T}), \\ \sigma_{T+1|T}^{2,RTV} &= \sigma_\eta^2 + \phi^2\sigma_v^2 = \sigma_\eta^2(1+\phi^2\delta), \end{aligned} \tag{15}$$

using results in Clements and Galvão (2013b) and Clements (2017).

A.2 | Noise revisions

The EOS and RTV forecasts of the mean are given by

$$\begin{aligned} \mu_{T+1|T}^{EOS} &= \phi y_T^{T+1} = \phi(y_T + \sigma_\varepsilon \xi_{3T}), \\ \mu_{T+1|T}^{RTV} &= B\phi y_T^{T+1} = B\phi(y_T + \sigma_\varepsilon \xi_{3T}). \end{aligned}$$

where (see Clements & Galvão, 2013b)

$$B = \frac{\sigma_y^2}{\sigma_y^2 + \sigma_\varepsilon^2} = \frac{\sigma_\eta^2/(1-\phi^2)}{\sigma_\eta^2/(1-\phi^2) + \sigma_\eta^2\delta} = \frac{(1-\phi^2)^{-1}}{((1-\phi^2)^{-1}+\delta)}.$$

The EOS and RTV forecasts of the variances are given by

$$\sigma_{T+1|T}^{2,EOS} = \sigma_\eta^2, \tag{16}$$

$$\sigma_{T+1|T}^{2,RTV} = \sigma_\eta^2(1+\delta+\varrho), \tag{17}$$

with $\varrho = [\phi^2(B-1)^2/(1-\phi^2) + \delta B^2\phi^2]$. Equation (17) is derived as

$$\begin{aligned} \sigma_{T+1|T}^{2,RTV} &= var(y_{T+1} + \sigma_\varepsilon \xi_{3T+1} - B\phi y_T - B\phi\sigma_\varepsilon \xi_{3T}) \\ &= var(y_t) + B^2\phi^2 var(y_t) + (1+B^2\phi^2)\sigma_\varepsilon^2 - 2B\phi Cov\,(y_t y_{t-1}) \\ &= \sigma_y^2(1+B^2\phi^2 - 2B\phi^2) + (1+B^2\phi^2)\sigma_\varepsilon^2 \\ &= \sigma_\eta^2(1+\delta+\varrho). \end{aligned}$$

Because $B < 1$ and $|\phi| < 1$, then $\varrho > 0$, implying that for the same δ and ϕ, $\sigma_{T+1|T}^{2,RTV}$ for noise is greater than $\sigma_{T+1|T}^{2,EOS}$ for news. If there are no revisions ($\delta = 0$), then $\varrho = 0$ since $B = 1$.

A.3 | Proposition 1

Proposition 1. *The difference between the EOS and RTV logscores (where the logscore is given by 11), $\Delta score^{News}$, for an AR(1) model, when the target is the initial release y_{T+1}^{T+2}, data revisions are pure news, and the DGP is given by (12), (13)* *and (14), is given by*

$$\Delta score^{News} = \frac{1}{2}\left[\frac{\delta(\phi^2-1)}{1+\delta} + \ln[(1+\delta)/(1+\phi^2\delta)]\right] \geq 0. \tag{18}$$

Proof. The difference between EOS and RTV log score when revisions are news:

$$\begin{aligned}\Delta score^{News} &= E[-\ln(p^{EOS}_{T+1|T}(y^{T+2}_{T+1}))] - E[-\ln(p^{RTV}_{T+1|T}(y^{T+2}_{T+1}))]\\ &= E[-\ln(p^{EOS}_{T+1|T}(y_{T+1}-\sigma_v\xi_{2T+1}))] - E[-\ln(p^{RTV}_{T+1|T}(y_{T+1}-\sigma_v\xi_{2T+1}))]\\ &= \left[\frac{1+\phi^2\delta}{2(1+\delta)} + \frac{1}{2}\ln(\sigma^2_\eta(1+\delta))\right] - \left[\frac{1}{2}+\frac{1}{2}\ln\sigma^2_\eta(1+\phi^2\delta)\right]\\ &= \frac{1}{2}\left[\frac{\delta(\phi^2-1)}{1+\delta} + \ln[(1+\delta)/(1+\phi^2\delta)]\right],\end{aligned}$$

and we need to show that $\Delta score^{News} \geq 0$ in order to establish the dominance of RTV over EOS on log score. If we take the derivative of the expression in brackets with respect to ϕ^2, we find $\delta/(1+\delta) - \delta/(1+\phi^2\delta)$. Because $(1+\delta) > (1+\phi^2\delta)$, since $\phi^2 < 1$ and $\delta \geq 0$, the derivative is always negative. This means that the minimum value of $\Delta score^{News}$ will be at the maximum value of ϕ^2, that is, $\phi^2 \approx 1$. Based on the expression above, it is clear that if $\phi^2 = 1$, $\Delta score^{News} = 0$. If $\Delta score^{News}$ is equal to zero at its minimum, then for values such that $0 \leq \phi^2 < 1$, we have $\Delta score^{News} \geq 0$. □

A.4 | Proposition 2

Proposition 2. *The difference between the EOS and RTV log scores, $\Delta score^{Noise}$, for an AR(1) model, when the target is the initial release y^{T+2}_{T+1}, data revisions are pure noise, and the DGP is given by (12), (13), and (14), is given by*

$$\Delta score^{Noise} = \frac{1}{2}\left[(\delta(1+\phi^2)) - \ln(1+\delta+\varrho)\right] \geq 0. \tag{19}$$

Proof. The difference between EOS and RTV log score when revisions are noise:

$$\begin{aligned}\Delta score^{Noise} &= E[-\ln(p^{EOS}_{T+1|T}(y^{T+2}_{T+1}))] - E[-\ln(p^{RTV}_{T+1|T}(y^{T+2}_{T+1}))]\\ &= E[-\ln(p^{EOS}_{T+1|T}(y_{T+1}+\sigma_\varepsilon\xi_{3T+1}))] - E[-\ln(p^{RTV}_{T+1|T}(y_{T+1}+\sigma_\varepsilon\xi_{3T+1}))]\\ &= \left[\frac{1}{2}(1+\delta(1+\phi^2)) + \frac{1}{2}\ln(\sigma^2_\eta)\right] - \left[\frac{1}{2}+\frac{1}{2}\ln(\sigma^2_\eta(1+\delta+\varrho))\right]\\ &= \frac{1}{2}\left[(\delta(1+\phi^2)) - \ln(1+\delta+\varrho)\right].\end{aligned}$$

To show that $\Delta score^{Noise} \geq 0$, we use the concavity of the logarithm function. But first note that we can rewrite ϱ as

$$\begin{aligned}\varrho &= [\phi^2(B-1)^2/(1-\phi^2) + \delta B^2\phi^2]\\ &= \phi^2(1-\phi^2)^{-1}\left[\frac{\delta^2}{((1-\phi^2)^{-1}+\delta)^2}\right] + \delta\phi^2\left[\frac{(1-\phi^2)^{-2}}{((1-\phi^2)^{-1}+\delta)^2}\right]\\ &= \frac{\phi^2(1-\phi^2)^{-1}\delta}{((1-\phi^2)^{-1}+\delta)} = \phi^2\delta B.\end{aligned}$$

Recall that $x \geq \ln(1+x)$ if $x \geq 0$. In the case that $B = 1$, $\varrho = \delta\phi^2$, and then $\delta(1+\phi^2) > \ln(1+\delta(1+\phi^2))$. When $\sigma^2_\varepsilon > 0$, then $B < 1$, and we have

$$\Delta score^{Noise} = \frac{1}{2}(\delta(1+\phi^2)) - \ln(1+\delta(1+\phi^2 B)).$$

Since $\phi^2 B < \phi^2$, then it must be the case that $\Delta score^{Noise} \geq 0$, establishing the dominance of RTV over EOS on log score for noise revisions. □