Real-time Forecasting of Inflation and Output Growth in the Presence of Data Revisions

Michael P. Clements and Ana Beatriz Galvão

No 953

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Real-time Forecasting of Inflation and Output Growth in the Presence of Data Revisions

Michael P. Clements
Department of Economics
University of Warwick
M.P.Clements@warwick.ac.uk

Ana Beatriz Galvão*
Department of Economics
Queen Mary University of London
a.ferreira@qmul.ac.uk

December 9, 2010

Abstract

We show how to improve the accuracy of real-time forecasts from models that include autoregressive terms by estimating the models on 'lightly-revised' data instead of using data from the latest-available vintage. Forecast accuracy is improved by reorganizing the data vintages employed in the estimation of the model in such a way that the vintages used in estimation are of a similar maturity to the data in the forecast loss function. The size of the expected reductions in mean squared error depend on the characteristics of the data revision process. Empirically, we find RMSFE gains of 2-4% when forecasting output growth and inflation with AR models, and gains of the order of 8% with ADL models.

Keywords: real-time data, news and noise revisions, optimal forecasts, multi-vintage models.

JEL code: C53.

*Corresponding author: Dr. Ana Beatriz Galvão. Queen Mary University of London, Department of Economics, Mile End Road, E1 4NS, London, UK. Phone: ++44-20-78828825. email: a.ferreira@qmul.ac.uk. We are grateful to Dick van Dijk and Simon van Norden for helpful comments, and to the participants at the 6th Forecasting Workshop at the ECB, Frankfurt, 2010, the 4th Time Series Conference, CIREQ, Montreal, 2010, and the 2010 Conference on Real-Time Data Analysis, Philadelphia.

1 Introduction

There has been much interest in the recent literature regarding the effects of different data vintages on model specification and forecast evaluation, and in the use of 'real-time data' in assessing predictability, as opposed to using 'final-revised' data, based on concerns that the use of final-revised data may exaggerate the predictive power of explanatory variables relative to what could actually have been achieved at the time using the then available data[1]. In this paper, we show how to improve the accuracy of real-time forecasts from models that include autoregressive terms by estimating them with lightly-revised data instead of using data from the latest-available vintage. We present a real-time analysis, in the sense that at each point in time the forecasting models are specified and the parameters estimated using only data for time periods up to that point in time, *and* the vintages of data used are restricted to those which would have been available at that time. Pseudo out-of-sample exercises adhere to the first aspect but use vintages of data that would not have been available at that time (as an example, see the study by Stock and Watson (2008)). A number of recent forecasting exercises have used only those vintages that would have been available in real time (see, e.g., Clements and Galvão (2008, 2009))

When computing forecasts in real-time, the majority of the literature uses the 'traditional approach' to real-time forecasting. At each point in time, the values of all the observations from the latest-available vintage of data are used to estimate the forecasting model. This is known as the end-of-sample vintage approach, or EOS, following Koenig, Dolmas and Piger (2003). To the extent that later estimates of a data point are more accurate or reliable than earlier estimates, this strategy uses the 'best' estimates of the data which are available at the time the forecast is made. However, it implies that a large part of the data used in model estimation has been revised many times, while the forecast is conditioned on data that has been just released or only revised a few times. In the context of autoregressive models, we show the traditional way of using real-time data for forecasting does not minimise the expected squared forecast error in population.

[1] See, for example, Diebold and Rudebusch (1991b, 1991a), Robertson and Tallman (1998), Orphanides (2001), Croushore and Stark (2001, 2003), Stark and Croushore (2002), Faust, Rogers and Wright (2003) and Orphanides and van Norden (2005).

Our main methodological contribution is to show that the use of 'real-time vintage' (RTV) can overcome the deficiencies of using EOS data.[2] Forecast accuracy is improved by reorganizing the data vintages employed in the estimation of the forecasting model in such a way that the data vintages used in model estimation are of a similar maturity to the vintage of data on which the forecast is conditioned. Even if the target is to forecast post-revision data, the approach that uses RTV data to estimate the forecasting model reduces mean squared error in comparison with the use of EOS data.

We compare the RTV approach to forecasting with a model that uses the multiple estimates of the same observation that are typically available.[3] It is not clear that using multiple data vintages would improve forecast accuracy. In a recent review of forecasting with real-time data, Croushore (2006) concludes that the results of forecasting with state-space models that incorporate data revisions are mixed, compared to simply ignoring data revisions. In this paper, we add to this knowledge base by presenting forecasts from a vector autoregression (VAR) that models the relationships between the different vintage estimates, in the spirit of recent work by Garratt *et al.* (2009, 2008) and Hecq and Jacobs (2009).

The approach of Kishor and Koenig (2010) (building on earlier contributions by Howrey (1978, 1984) and Sargent (1989)) nicely contrasts ours. Because the forecast will be conditioned on lightly-revised data, we estimate the model in such a way that the parameters are optimal for the generation of forecasts which are conditioned on lightly-revised data. Kishor and Koenig (2010) instead solve the problem by estimating the model on (largely) post-revision data, which necessarily means using data that stops short of the forecast origin. The model forecasts of the periods up to the origin are combined with lightly-revised data for these periods via the Kalman filter to obtain post-revision estimates of these latest data points. The estimated model is then applied to these data estimates to generate forecasts of revised values of future observations. Hence they essentially estimate the

[2] Harrison, Kapetanios and Yates (2005) consider the use of EOS data when there are measurement errors, and suggest that the most recent observations might be downweighted.

[3] Examples of multiple-vintage models include Harvey, McKenzie, Blake and Desai (1983), Howrey (1984), Patterson (1995, 2003), Jacobs and van Norden (2007), Cunningham, Eklund, Jeffery, Kapetanios and Labhard (2009), Garratt, Lee, Mise and Shields (2009, 2008) and Hecq and Jacobs (2009).

model on revised data, and condition the forecasts on revised data. Our approach estimates the model, and conditions the forecasts, on lightly-revised data. Our approach is simple OLS and does not require any kind of filtering, and the forecast target can be either first-released or post-revision data.

We derive analytical results that relate the properties of the forecasts generated by the use of EOS and RTV data to the properties of data revisions for an assumed data generation process that characterises data revisions as ‘news’ or ‘noise’, in the Mankiw and Shapiro (1986) sense. We use the Jacobs and van Norden (2007) statistical framework to model the ‘regular’ rounds of revisions which are made to the data at a level of detail that allows us to delineate between first and subsequent revisions, as the sizes of the variances of the first and subsequent revisions are found to affect the relative performance of RTV and EOS. It might be argued that data revisions are irregular and not amenable to modelling as a stationary process. This is likely to be true of benchmark revisions but the evidence presented by Croushore (2006) suggests that growth rates - the focus of our analysis - will be affected to a lesser degree than the levels of variables.[4] Our data generating process neglects certain characteristics of US data revisions - such as the seasonal nature of some revisions, as we explain below - but such complications would not affect the finding that RTV improves forecast accuracy relative to EOS in population.

For AR models we are able to relate the population properties of the estimators (using RTV or EOS data) back to the properties of the hypothesized data generation process, including the properties of the revisions process, which yields additional insights when we can clearly categorize a series in terms of news or noise revisions. In models with explanatory variables it will be more difficult to obtain the direction of any bias of the estimator without making a range of further assumptions, including specifying the covariances between the revisions to the series being forecast and the revisions to the explanatory variables. Hence our analytical results are for the AR, as this

[4]Siklos (2008) identifies eight benchmark revisions in 1966, 1971, 1976, 1981, 1986, 1992, 1996 and 2001, all occurring in the data vintage of the first quarter of the year - so, for example, the 1981:1 data set has data up to 1980:4 calculated on a different basis or definition to the 1980:4 vintage data set. The way which the national accounts data are calculated then remains unchanged until the 1986:1 data set. Base year changes occurred in 1976, 1985 and 1991.

allows for sharper predictions, although the general arguments based on the nature of the real-time forecast loss function that support the use of RTV data in the case of AR models are also applicable to ADL models.[5]

Although our population analytical results show that it is always better to use RTV data instead of EOS data, of interest is whether gains will be observed once an allowance is made for parameter estimation uncertainty. Hence we extend our results with a Monte Carlo. The results suggest that the gains to using RTV data relative to EOS data are likely to be modest, of the order of around 2-3% on mean squared error for reasonably large samples. For small samples, larger gains of around 3-8% might occur in the case of revisions which add news. These results are based on what might be regarded as reasonably typical, empirically-calibrated data generation processes. We also check the robustness of our finding that RTV tends to dominate EOS in small samples by simulating data from an estimated vintage-VAR model, and carrying out the same forecasting exercise. The VAR model is agnostic as to the nature of data revisions. In essence the Monte Carlo results are unchanged, indicating that our findings are unaffected by the precise way in which the data and data revisions are modelled.

Nevertheless, all the Monte Carlo results we report hold fixed factors that might be relevant empirically. Chief amongst these are the assumed constancy of the underlying models of the variables and their revisions over time. Of interest will be the usefulness of these results as a guide to empirical outcomes. In our empirical forecasting exercises we compare the use of EOS and RTV data as competing approaches to real-time forecasting, and look at the extent to which the empirical findings are consistent with our analysis. Our analytical results are directly applicable to AR models, but we also consider ADL models of output growth and inflation.

The plan of the rest of the paper is as follows. Section 2 describes the real-time forecasting setting, and provides some intuition as to why simply using the latest-available vintage data to

[5]Koenig *et al.* (2003) derive expressions for the bias of the estimator of distributed lag (DL) models for various ways of using real-time vintage data, but do not explicitly model the revisions process. They simply specify the first and final estimates, and hence have a single revision. We model the sequence of revisions, and as noted, find that the properties of the estimators depend on the relative variances of the successive rounds of revision.

estimate the model may not be the best strategy. Section 3 presents the statistical framework that we use to model data revisions. This is essentially the Jacobs and van Norden (2007) state-space model, but modified to allow for non-zero mean revisions, because this is a feature of some US macroeconomic aggregates we consider. Section 4 presents one of the main results of the paper, that the traditional way of using real-time data (EOS) to estimate autoregressive models generates forecasts that are not optimal, and are biased when data revisions are non-zero mean. Section 5 shows how the use of RTV data to estimate the autoregressive model generates optimal, unbiased forecasts. Section 6 presents the Monte Carlo investigation of the small-sample relevance of our analytical results. This uses the state-space model of section 3 to generate data subject to noise and news revisions, as well as a VAR model, which is also briefly described. Section 7 is the empirical forecast comparison. Section 8 offers some concluding remarks. The derivations of the main results are confined to an appendix.

2 Motivation

When we allow that data are subject to revision, the forecaster at period $T+1$ will have access to the vintage $T+1$ values of the observations on y up to time period T, as the first-release is published with a one quarter lag. We let y_t^{t+j} denote the vintage $t+j$ estimate of the value of the variable in period t, where $j = 1, 2, 3, \ldots$, and where $j = 1$ denotes the first-release value. Hence the forecaster has $(y_1^{T+1}, \ldots, y_{T-1}^{T+1}, y_T^{T+1})$. This is the 'latest-available vintage' data at period $T+1$, which we can write as $\left\{y_i^{T+1}\right\}_{i=1,2,\ldots,T.}$. But the forecaster will also have the previous vintages, for example, the $T+1-j$ vintage, $\left\{y_i^{T+1-j}\right\}$ for $j = 1, 2, 3, \ldots$, and where $i = 1, 2, \ldots, T-j$.

The traditional approach estimates the forecasting model on the latest-available $(T+1)$ vintage, and conditions the forecasts on the $T+1$ vintage values of the forecast-origin data. So for an AR(2), the model is estimated on:

$$y_t^{T+1} = \alpha_0 + \alpha_1 y_{t-1}^{T+1} + \alpha_2 y_{t-2}^{T+1} + e_{t,EOS} \tag{1}$$

for $t = 3, \ldots, T$, and the forecast of y_{T+1} is given by:

$$\widehat{y_{T+1}}_{,EOS} = \hat{\alpha}_0 + \hat{\alpha}_1 y_T^{T+1} + \hat{\alpha}_2 y_{T-1}^{T+1}.$$

The parameter estimates $(\hat{\alpha}_0, \hat{\alpha}_1\, \hat{\alpha}_2)$ are based on data that are for the most part post-revision (or heavily revised) data. Suppose for the sake of argument that data are revised 14 times, so that $y_t^{t+15} = \widetilde{y}_t$, where $\widetilde{y}_t$ is the true value, then a proportion $(T-14)/T$ of the T estimation observations underlying (1) are fully-revised or true data. Yet the forecasts are conditioned on a first-release observation and a once-revised observation (y_T^{T+1} and y_{T-1}^{T+1} respectively). This is the 'apples and oranges' problem of Kishor and Koenig (2010).

Given that the forecast will be conditioned on early-release data (y_T^{T+1} and y_{T-1}^{T+1}), in the case of an AR(2), the RTV approach estimates the parameters of the AR(2) on matching early-release data:

$$y_t^{t+1} = \beta_0 + \beta_1 y_{t-1}^t + \beta_2 y_{t-2}^t + e_{t,RTV}$$

for $t = 3, \ldots, T$, and the forecast of y_{T+1} is given by:

$$\widehat{y_{T+1}}_{,RTV} = \hat{\beta}_0 + \hat{\beta}_1 y_T^{T+1} + \hat{\beta}_2 y_{T-1}^{T+1}.$$

The forecast is conditioned on exactly the same data as under the traditional approach - the latest values from the most recent vintage. We show in what follows that it is not optimal to use predominantly heavily-revised data from the most recent vintage to estimate the model. Notice that the RTV approach uses first-release data for the first lag, and second-release data for the second lag, and that this generalizes for an AR(p), so that it is not the case that only first-release data is used to estimate the model. The important point is that the data maturities used in estimation match those that the forecasts are conditioned on.

In what follows (section 4 and 5) we show analytically that our RTV approach yields more accurate forecasts. To do that, we need a data generating process for the different data vintages and the true values.

3 Statistical framework

Generally, the basic statistical framework for modelling data revisions relates a data vintage estimate to the true value plus an error or errors, where the errors are typically unobserved. So the period $t+s$ vintage estimate of the value of y in period t, denoted y_t^{t+s}, where $s = 1, \ldots, l$, consists of the true value $\widetilde{y}_t$, as well as (in the general case) news and noise components, v_t^{t+s} and ε_t^{t+s}, so that $y_t^{t+s} = \widetilde{y}_t + v_t^{t+s} + \varepsilon_t^{t+s}$. Data revisions are news when initially released data are optimal forecasts of later data, so news revisions are not correlated with the earlier-release data, $Cov\left(v_t^{t+s}, y_t^{t+s}\right) = 0$. Data revisions are noise when each new release of the data is equal to the true value of y_t, denoted $\widetilde{y}_t$, plus noise, so that noise revisions are not correlated with the truth, $Cov\left(\varepsilon_t^{t+s}, \widetilde{y}_t\right) = 0$. We adopt the framework of Jacobs and van Norden (2007) which stacks the l different vintage estimates of y_t, namely, $y_t^{t+1}, \ldots, y_t^{t+l}$ in the vector $\mathbf{y}_t = \left(y_t^{t+1}, \ldots, y_t^{t+l}\right)'$, and similarly $\boldsymbol{\varepsilon}_t = \left(\varepsilon_t^{t+1}, \ldots, \varepsilon_t^{t+l}\right)'$ and $\mathbf{v}_t = \left(v_t^{t+1}, \ldots, v_t^{t+l}\right)'$, so that:

$$\mathbf{y}_t = \mathbf{i}\widetilde{y}_t + \mathbf{v}_t + \boldsymbol{\varepsilon}_t \tag{2}$$

where $\mathbf{i}$ is a l-vector of ones. One way of defining a revisions process with the required characteristics is to assume a process for $\widetilde{y}_t$, for example, an AR(p) with iid disturbances $R_1\eta_{1t}$, plus a sum of l news components $v_{i,t}$:

$$\widetilde{y}_t = \rho_0 + \sum_{i=1}^{p} \rho_i \widetilde{y}_{t-i} + R_1\eta_{1t} + \sum_{i=1}^{l} v_{i,t}, \tag{3}$$

where $v_{i,t} = \sigma_{v_i}\eta_{2t,i}$ (for $i = 1, ..., l$) and both η_{1t} and $\eta_{2t,i}$ are $iid(0,1)$. We let $\rho(L) = \sum_{i=1}^{p} \rho_i L^i$ and assume that the roots of $(1 - \rho(L)) = 0$ lie outside the unit circle, so that $\widetilde{y}_t$ is a stationary process. The news and noise components of each vintage in $\mathbf{y}_t$ are:

$$\mathbf{v}_t = \begin{bmatrix} v_t^{t+1} \\ v_t^{t+2} \\ \vdots \\ v_t^{t+l} \end{bmatrix} = - \begin{bmatrix} \sum_{i=1}^{l} v_{i,t} \\ \sum_{i=2}^{l} v_{i,t} \\ \vdots \\ v_{l,t} \end{bmatrix}, \quad \boldsymbol{\varepsilon}_t = \begin{bmatrix} \varepsilon_t^{t+1} \\ \varepsilon_t^{t+2} \\ \vdots \\ \varepsilon_t^{t+l} \end{bmatrix} = \begin{bmatrix} \sigma_{\varepsilon_1}\eta_{3t,1} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \vdots \\ \sigma_{\varepsilon_l}\eta_{3t,l} \end{bmatrix}, \tag{4}$$

where $\eta_{3t,i}$ is $iid(0,1)$. The shocks are also mutually independent, that is, if $\boldsymbol{\eta}_t = [\eta_{1t}, \boldsymbol{\eta}'_{2t}, \boldsymbol{\eta}'_{3t}]$, then $E(\boldsymbol{\eta}_t) = 0$, with $E(\boldsymbol{\eta}_t\boldsymbol{\eta}'_t) = I$.

Therefore, the first estimate of y_t, y_t^{t+1}, is $y_t^{t+1} = \rho_0 + \sum_{i=1}^{p} \rho_i \widetilde{y}_{t-i} + R_1\eta_{1t} + \sigma_{\varepsilon_1}\eta_{3t,1}$, which does not include any news component. Later estimates may be characterised by noise, but include more news components. For example, $y_t^{t+4} = \rho_0 + \sum_{i=1}^{p} \rho_i \widetilde{y}_{t-i} + R_1\eta_{1t} + \sigma_{\varepsilon_4}\eta_{3t,4} + \sum_{i=1}^{3} v_{i,t}$ is a more accurate estimate of $\widetilde{y}_t$ than y_t^{t+1}, because it includes the news terms $(\sum_{i=1}^{3} v_{i,t})$ which are part of $\widetilde{y}_t$ (in addition to the noise component). As noted by Mankiw and Shapiro (1986), news revisions imply that $var(y_t^{t+1}) < var(y_t^{t+l})$, while noise revisions imply that $var(y_t^{t+1}) > var(y_t^{t+l})$, assuming that later estimates are less 'noisy' ($\sigma_{\varepsilon_1} > \sigma_{\varepsilon_l}$). If $\sigma_{v_l} = 0$ and $\sigma_{\varepsilon_l} = 0$ the l-vintage value is the true value, $y_t^{t+l} = \widetilde{y}_t$. The assumption that $\widetilde{y}_t$ is a stationary process ensures that $\mathbf{y}_t$ is a stationary process from (2), as both the news and noise terms are stationary.

The assumptions we have made imply that both noise and news revisions are zero mean, so that the unconditional mean of the underlying series $\{\widetilde{y}_t\}$ and the observed data $\{\mathbf{y}_t\}$ are equal at $\rho_0 (1 - \rho(1))^{-1}$. However, there is evidence that the revisions to some macroeconomic data are non-zero mean, as we shall discuss in section 6.[6] To account for this characteristic of the revisions process, we consider a modified version of the statistical model that allows data revisions to affect the mean of the $\{\widetilde{y}_t, \mathbf{y}_t\}$. We assume that each news term is instead $v_{i,t} = \mu_{v_i} + \sigma_{v_i}\eta_{2t,i}$, and the noise component is $\varepsilon_t^{t+i} = -\mu_{\varepsilon_i} + \sigma_{\varepsilon_i}\eta_{3t,i}$. The true process is:

[6] See Aruoba (2008) and Corradi, Fernandez and Swanson (2009) for recent analyses of the properties of data revisions.

$$\widetilde{y}_t = \left[\rho_0 + \sum_{i=1}^{l} \mu_{v_i}\right] + \sum_{i=1}^{p} \rho_i \widetilde{y}_{t-i} + R_1 \eta_{1t} + \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i}, \tag{5}$$

since now $\sum_{i=1}^{l} v_{it} = \sum_{i=1}^{l} \mu_{v_i} + \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i}$. The news and noise processes of each vintage are:

$$\mathbf{v}_t = -\begin{bmatrix} \sum_{i=1}^{l} \mu_{v_i} \\ \sum_{i=2}^{l} \mu_{v_i} \\ \vdots \\ \mu_{v_l} \end{bmatrix} - \begin{bmatrix} \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i} \\ \sum_{i=2}^{l} \sigma_{v_i} \eta_{2t,i} \\ \vdots \\ \sigma_{v_l} \eta_{2t,l} \end{bmatrix}, \ \boldsymbol{\varepsilon}_t = -\begin{bmatrix} \mu_{\varepsilon_1} \\ \mu_{\varepsilon_2} \\ \vdots \\ \mu_{\varepsilon_l} \end{bmatrix} + \begin{bmatrix} \sigma_{\varepsilon_{1,t}} \eta_{3t,1} \\ \sigma_{\varepsilon_{2,t}} \eta_{3t,2} \\ \vdots \\ \sigma_{\varepsilon_{l,t}} \eta_{3t,l} \end{bmatrix}. \tag{6}$$

The statistical model can be cast in state-space form using (2) as the observation equation and combining (5) and (6) to obtain the measurement equation. The parameters can be estimated by maximum likelihood using the Kalman Filter, as described by Jacobs and van Norden (2007).

The formulation of the model that we have described in this section assumes that the revisions process is the same irrespective of the quarter of the year to which time period t belongs. However this is a simplification of the way the Bureau of Economic Analysis (BEA, the US statistical agency) operates. The BEA releases an 'advance', 'preliminary' and 'final' estimate of real GDP growth at about one month, two months, and three months after the end of the quarter. We observe the 'advance' estimate, which we call the first release (y_t^{t+1}), and then the next quarter we observe the 'final' estimate (y_t^{t+2}). The data are then unchanged until the July revision, so that whether $y_t^{t+3} - y_t^{t+2}$ is non-zero will depend on the quarter of the year to which t belongs.[7] In appendix B, we outline a 'seasonal' version of the model described in this section which allows for this feature of the revisions process. However, the key analytical findings in the following sections would remain unchanged. The second moment matrices in the expressions for the estimators given in the following two sections would become more complicated, but the qualitative properties of the estimators remain as stated.

[7] The data are subject to three annual revisions which occur in the July of each year, as described by, e.g., Fixler and Grimm (2005, 2008) and Landefeld, Seskin and Fraumeni (2008). This suggests that $\{y_t^{t+14}\}$ will have undergone all the regular revisions irrespective of which quarter of the year t falls in.

4 Estimating and Forecasting with AR models using EOS data

In this section we derive the values of the AR(p) parameters that minimise the real-time expected squared forecast error assuming the model of data revisions described in the last section. For ease of exposition, we first assume that revisions are zero mean. We then show that the traditional way of using real-time data for forecasting does not generate forecasts that minimise the real-time loss function. Finally, we assess the impact of non-zero mean revisions on the traditional way of forecasting with AR models in real-time.

4.1 Optimal AR parameters in population

Consider how forecasters normally use real-time data to compute forecasts with an autoregressive model. At time $T+1$, the $T+1$ vintage of data contains data up to period T, so that the $T+1$ vintage is used to estimate the AR(p) model, and the forecasts are obtained by conditioning on the model estimates and the lagged values of y from the latest data vintage, namely, $\mathbf{y}_T^{T+1} = \left(y_T^{T+1}, \ldots, y_{T-p+1}^{T+1}\right)'$. But does least squares estimation of the AR model using the $T+1$ data vintage minimise the expected squared forecast error? To answer this question, we first obtain the population value of the parameter vector that minimizes the real-time squared forecast error for a forecast conditioned on $\mathbf{y}_T^{T+1}$, when the data are subject to revisions as described in section 3. For a p^{th}-order autoregression, the forecast is given by the combination $\phi_0 + \boldsymbol{\phi}'\mathbf{y}_T^{T+1}$, where ϕ_0 is the intercept and $\boldsymbol{\phi}' = \left(\phi_1, \ldots, \phi_p\right)$ contains the slope parameters. The optimal parameter values $(\phi_0^*, \boldsymbol{\phi}^*)$, in terms of minimizing the real-time squared-error loss, are the solution to:

$$(\phi_0^*, \boldsymbol{\phi}^*) = \underset{\phi_0, \boldsymbol{\phi}}{\arg\min} E\left(L\left(\phi_0, \boldsymbol{\phi}\right)\right)$$

where:

$$L\left(\phi_0, \boldsymbol{\phi}\right) = \left(y_{T+1}^{T+1+f} - \phi_0 - \boldsymbol{\phi}'\mathbf{y}_T^{T+1}\right)^2. \tag{7}$$

In (7), the vintage of data employed to compute the forecast errors is that available at $T+1+f$, so that $f=1$ indicates first-release data is used. The following proposition is derived in the Appendix.

Proposition 1 *The optimal parameters $(\phi_0^*, \boldsymbol{\phi}^*)$ when the data generating process is given by (2)–(4), are*

$$\begin{aligned}
\boldsymbol{\phi}^* &= \left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \Sigma_{\mathbf{v}} + \Sigma_{\widetilde{\mathbf{y}}\mathbf{v}} + \Sigma'_{\widetilde{\mathbf{y}}\mathbf{v}} + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right)^{-1}\left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \boldsymbol{\Sigma}'_{\widetilde{\mathbf{y}}\mathbf{v}}\right)\boldsymbol{\rho} \\
\phi_0^* &= \left(1 - \boldsymbol{\phi}^{*\prime}\mathbf{i}\right)\mu_{\widetilde{y}}. \qquad (8)
\end{aligned}$$

The slope parameter simplifies to $\boldsymbol{\phi}^*_{news} = \left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \Sigma_{\mathbf{v}} + \Sigma_{\widetilde{\mathbf{y}}\mathbf{v}} + \Sigma'_{\widetilde{\mathbf{y}}\mathbf{v}}\right)^{-1}\left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \boldsymbol{\Sigma}'_{\widetilde{\mathbf{y}}\mathbf{v}}\right)\boldsymbol{\rho}$, if data revisions add news, and is $\phi^*_{noise} = \left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right)^{-1}\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}}\boldsymbol{\rho}$ if data revisions only reduce noise, where the second moment matrices $\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}}$, $\Sigma_{\mathbf{v}}$, $\Sigma_{\widetilde{\mathbf{y}}\mathbf{v}}$, and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ are defined in the Appendix, $\mathbf{i}$ is a p-vector of 1's, and $\mu_{\widetilde{y}} \equiv E(\widetilde{y}_t)$. The optimal values hold for all $f \geq 1$, i.e., irrespective of whether the goal is to forecast the first-released value (y_{T+1}^{T+2}) or the latest available estimate (y_{T+1}^{T+l}).

For the special case of an AR(1) model, for general revisions that are a combination of news and noise:

$$\phi_1^* = \frac{\rho_1\left(\sigma^2_{\widetilde{y}} - \sigma^2_v\right)}{\sigma^2_{\widetilde{y}} - \sigma^2_v + \sigma^2_{\varepsilon_1}}$$

where $\sigma^2_v \equiv \sum_{i=1}^{l}\sigma^2_{v_i}$, and $\sigma^2_{\widetilde{y}} = Var(\widetilde{y}_t)$. As a consequence, for pure news $(\sigma^2_{\varepsilon_1} = 0)$:

$$\phi^*_{1,news} = \rho_1, \quad \phi^*_{0,news} = \rho_0.$$

Note that $\phi^*_{1,news} = \rho_1$ only holds for $p = 1$: in general when there are news revisions the parameter vector of the underlying process $\widetilde{y}_t$ (i.e., $\boldsymbol{\rho}$) is not optimal from a forecasting perspective when the forecasts are conditioned on early estimates, as is typically the case in a real-time forecasting exercise.

For pure noise $(\sigma^2_v = 0)$:

$$\phi^*_{1,noise} = \frac{\rho_1\sigma^2_{\widetilde{y}}}{\sigma^2_{\widetilde{y}} + \sigma^2_{\varepsilon_1}} \qquad (9)$$

so that $\phi^*_{1,noise} < \rho_1$ provided $\sigma^2_{\varepsilon_1} \neq 0$.

4.2 Estimating AR forecasting models using EOS data

As noted, when forecasting with AR models using real-time data, the standard approach is to replace the model estimation data with the latest estimates of all the past observations given in the data vintage available at the forecast origin. So, for example, at time $T+1$, the $T+1$ vintage of data contains data up to T, and is used for estimation, while at $T+2$, the $T+2$ vintage is used for estimation. We call this use of real-time data '*end-of-sample*' vintage data (EOS).

For forecasting y^{T+1+f}_{T+1} the AR forecasting model is given by:

$$y^{T+1}_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i y^{T+1}_{t-i} + e_{t,EOS} \tag{10}$$

where $t = p+1, \ldots, T$, assuming the latest data vintage is dated $T+1$. In matrix notation:

$$Y^{T+1} = \mathbf{i}\alpha_0 + \mathbf{Y}_{-1}\boldsymbol{\alpha} + error$$

where $\mathbf{Y}_{-1} = \left[Y^{T+1}_{-1}, \ldots, Y^{T+1}_{-p}\right]$, $\mathbf{i}$ is a $T-p$ vectors of 1's, and the vectors of observations Y^{T+1} and Y^{T+1}_{-i}, $i = 1, \ldots, p$, are:

$$Y^{T+1} = \left[y^{T+1}_{p+1}, \ldots, y^{T+1}_{T-1}, y^{T+1}_{T}\right]', \quad Y^{T+1}_{-i} = \left[y^{T+1}_{p+1-i}, \ldots, y^{T+1}_{T-i-1}, y^{T+1}_{T-i}\right]'$$

for $i = 1, \ldots, p$. Notice that the more recent y's will therefore have been revised fewer than l times.

The main result is summarized in the following proposition, which is derived in the appendix.

Proposition 2 *The population (asymptotic) value of the least-squares estimator of the parameter vector in the autoregressive model using EOS data, when the data are generated by (2)–(4), are* *given by:*

$$
\begin{aligned}
\boldsymbol{\alpha}^* &= \left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \boldsymbol{\Sigma}_{\underline{\mathbf{v}}} + \boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}\underline{\mathbf{v}}} + \boldsymbol{\Sigma}'_{\widetilde{\mathbf{y}}\underline{\mathbf{v}}} + \boldsymbol{\Sigma}_{\underline{\boldsymbol{\varepsilon}}}\right)^{-1} \left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}} + \boldsymbol{\Sigma}'_{\widetilde{\mathbf{y}}\underline{\mathbf{v}}}\right) \boldsymbol{\rho} \\
\alpha_0^* &= \left(1 - \boldsymbol{\alpha}^{*\prime}\mathbf{i}\right) \mu_{\widetilde{y}},
\end{aligned} \tag{11}
$$

where $\Sigma_{\underline{\mathbf{v}}}$ and $\Sigma_{\underline{\boldsymbol{\varepsilon}}}$ are second moment matrices of the news and noise components, and $\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}\underline{\mathbf{v}}}$ is the second moment matrix between the news and the underlying process, $\widetilde{y}_t$, and $\mu_{\widetilde{y}} \equiv E\left(\widetilde{y}_t\right)$.

A comparison of (11) and (8) shows that the conventional use of real-time data (EOS) for estimation of the AR(p) model does not deliver the optimal population parameters when there are data revisions. That is,

$$\boldsymbol{\alpha}^* \neq \boldsymbol{\phi}^* \text{ and } \alpha_0^* \neq \phi_0^*,$$

so that the forecasts of $\tau + 1$ computed using $\alpha_0^* + \boldsymbol{\alpha}^{*\prime}\mathbf{y}_\tau^{\tau+1}$ (for a set of forecast origins $\tau = T, T+1, T+2, \ldots$), where recall that $\mathbf{y}_\tau^{\tau+1} = \left(y_\tau^{\tau+1}, \ldots, y_{\tau-p+1}^{\tau+1}\right)'$, are not optimal in a squared-error loss sense. Intuitively, when the sample is large, the use of EOS data amounts to mainly using fully-revised data (i.e., data from the y_t^{t+l} vintage) whilst optimal forecasts are obtained by relating the first estimates of the LHS variable to early estimates of the RHS variables. The finding of the lack of optimality of EOS forecasts holds for news and noise revisions, although forecasts are unbiased when the data revisions are zero mean, as described by the following remark, derived in the Appendix.

Remark 1 *Forecasts computed using EOS data using the AR model with $(\alpha_0^*, \boldsymbol{\alpha}^*)$ are unbiased when data revisions are described by (2) – (4).*

Consider the special case of an AR(1). When revisions are news, we can show that the EOS estimator simplifies such that $\alpha_1^* = \rho_1$, matching the optimal value, but this is true only for the special case of $p = 1$.

Under noise:

$$\alpha_1^* = \frac{\rho_1 \sigma_{\widetilde{y}}^2}{\sigma_{\widetilde{y}}^2 + \sigma_{\varepsilon_l}^2} \tag{12}$$

An immediate implication is that $|\alpha_1^*| > |\phi^*|$ if earlier revisions are larger than later revisions (compare (12) to (9) when $\sigma_{\varepsilon_1}^2 > \sigma_{\varepsilon_l}^2$). Note that if $\sigma_{\varepsilon_l}^2 = 0$, so that the truth is eventually revealed when there is noise, then $\alpha_1^* = \rho_1$ for a large estimation sample. Even so, ρ_1 is not the parameter vector that minimizes the real-time squared forecast loss ($\phi_1^* \neq \rho_1$).

4.3 Optimal AR coefficients in population when revisions are non-zero mean

In this section we allow for non-zero mean revisions. Non-zero mean revisions indicate that in general $E\left(y_t^{t+r}\right) \neq E\left(y_t^{t+s}\right)$, so that forecasts may be biased for some vintage estimates of the actuals but not for others. We need to take a stance on the vintage, i.e., f, in y_{T+1}^{T+1+f}. Suppose the objective is to predict the first-released value, namely, y_{T+1}^{T+2}, so the optimal values $(\phi_0^*, \boldsymbol{\phi}^*)$ are defined as the solution of (7) with $f = 1$.

Proposition 3 *The optimal parameters $(\phi_0^*, \boldsymbol{\phi}^*)$ when the data are generated by (2), (5) and (6), for the loss function given by (7) with $f = 1$, are equal to $\boldsymbol{\phi}^*$, as derived in Proposition 1, and*

$$\phi_0^* = \left(1 - \boldsymbol{\phi}^{*\prime}\mathbf{i}\right)\mu_{\widetilde{y}} - \textstyle\sum_{i=1}^{l} \mu_{v_i} - \mu_{\varepsilon_1} - \boldsymbol{\phi}^{*\prime}\boldsymbol{\mu}_{\varepsilon} - \boldsymbol{\phi}^{*\prime}\boldsymbol{\mu}_{v}, \tag{13}$$

where $\mu_{\widetilde{y}} = (1 - \rho(1))^{-1}\left[\rho_0 + \sum_{i=1}^{l} \mu_{v_i}\right]$, $E\left(\varepsilon_t^{t+1}\right) = -\mu_{\varepsilon_1}$, $E\left(v_t^{t+1}\right) = -\sum_{i=1}^{l} \mu_{v_i}$, and $\boldsymbol{\mu}_{\varepsilon} = E(\boldsymbol{\varepsilon}_t)$ and $\boldsymbol{\mu}_{v} = E(\mathbf{v}_t)$ are $p \times 1$ vectors of the means of noise and news components.

4.4 Estimating AR forecasting models using EOS data when revisions are non-zero mean

In the Appendix we derive the following proposition:

Proposition 4 *The population (asymptotic) value of the least-squares estimator of the parameter vector in the autoregressive model using EOS data (eq. (10)), when the data are generated by (2),* *(5) and (6), is equal to the slope value of* $\boldsymbol{\alpha}^*$ *given in Proposition 2, and the intercept is given by*

$$\alpha_0^* = \left(1 - \boldsymbol{\alpha}^{*\prime}\mathbf{i}\right)\left(\mu_{\widetilde{y}} - \mu_{v_l} - \mu_{\varepsilon_l}\right), \tag{14}$$

where $\mu_{\widetilde{y}} = (1 - \rho(1))^{-1}\left[\rho_0 + \sum_{i=1}^{l}\mu_{v_i}\right]$, $E\left(\varepsilon_t^{t+l}\right) = -\ \mu_{\varepsilon_l}$ *and* $E\left(v_t^{t+l}\right) = -\mu_{v_l}$.

Comparing (13) with (14), we have that $\alpha_0^* \neq \phi_0^*$, that is, the EOS estimation will in general yield biased forecasts when revisions are non-zero mean.

Remark 2 *Forecasts of the first-released value computed using the AR model with parameter vector* $(\alpha_0^*, \boldsymbol{\alpha}^*)$*, as under EOS, are biased when data revisions are described by (2), (5) and (6), with bias of* $\left(\mu_{v_l} - \sum_{i=1}^{l}\mu_{v_i}\right) + \left(\mu_{\varepsilon_l} - \mu_{\varepsilon_1}\right) - \boldsymbol{\alpha}^{*\prime}\left[\left(\boldsymbol{\mu}_v + \mathbf{i}\mu_{v_l}\right) + \left(\boldsymbol{\mu}_\varepsilon + \mathbf{i}\mu_{\varepsilon_l}\right)\right]$.

Summarizing, the use of EOS data to estimate AR models for forecasting in real-time when data are subject to news and noise revisions delivers predictions that are not optimal, that is, the resulting forecasts do not minimise expected quadratic forecast loss. We show in the following section that alternative forecasts conditioned on the same information set deliver a smaller squared-error loss in population. When in addition data revisions are non-zero mean, forecasts using EOS are generally biased for the first-release of the target variable, whereas unbiased forecasts are easily obtained from the AR model by organizing the data as described in the next section.

5 Estimating AR forecasting models using RTV data

In this section we consider a way of using real-time data, motivated by the approach suggested by Koenig *et al.* (2003) in the context of distributed lag models, that delivers optimal estimators of the forecasting model in population. We adapt their approach to the estimation of AR models by regressing the period $t+1$ vintage value of y_t on the t-vintage data values of the lags, y_{t-i}, $i = 1, \ldots, p$, for $t = p+1, \ldots, T$. The reason for using the t-vintage value of the lags (rather than the $t+1$-vintage value) is that when the model is used for forecasting, say, the $T+2$ vintage value of y_{T+1}, the forecast will be conditioned on the previous $T+1$-vintage values of the explanatory variables. This use of vintage data is called '*real-time-vintage*' data (RTV). The model is:

$$y_t^{t+1} = \beta_0 + \sum_{i=1}^{p} \beta_i y_{t-i}^t + e_{t,RTV} \tag{15}$$

where $t = p+1, \ldots, T$. In matrix notation:

$$Y^t = \mathbf{i}\beta_0 + \mathbf{Y}_{-1}^t \boldsymbol{\beta} + error$$

where Y^t and$\mathbf{Y}_{-1}^t = \left[Y_{-1}^t, \ldots, Y_{-p}^t\right]$ are given by:

$$Y^t = \left[y_{p+1}^{p+2}, \ldots, y_{T-1}^T, y_T^{T+1}\right]', \quad Y_{-i}^t = \left[y_{p+1-i}^{p+1}, \ldots, y_{T-i-1}^{T-1}, y_{T-i}^T\right]', \quad i = 1, \ldots, p.$$

Note that Y^t and Y_{-i}^t contain early vintages of data relative to the period $T+1$-vintage: these observations are not replaced with the latest available ($T+1$-vintage) values for these observations.

Consider estimating equation (15) by OLS. A typical observation on the LHS and RHS variables is $\left\{y_t^{t+1}, \mathbf{y}_{t-1}^t = \left(y_{t-1}^t \ldots y_{t-p}^t\right)'\right\}$. This is a covariance stationary process, so we can calculate the population values of the OLS estimators as the values that satisfy:

$$(\beta_0^*, \boldsymbol{\beta}^*) = \underset{\beta_0, \boldsymbol{\beta}}{\arg\min} E\left(y_t^{t+1} - \beta_0 - \boldsymbol{\beta}\mathbf{y}_{t-1}^t\right)^2.$$

This estimation loss function is identical to the real-time forecast loss function (7), when $f = 1$, so that the solutions to the two coincide. Hence the use of RTV data to estimate the AR model delivers optimal forecasts, that is,

$$\beta_0^* = \boldsymbol{\phi}^*; \ \boldsymbol{\beta}^* = \phi_0^*.$$

The unbiasedness of using RTV data to compute forecasts follows directly because $(\beta_0^*, \boldsymbol{\beta}^*)$ satisfy $E\left(y_t^{t+1} - \beta_0 - \boldsymbol{\beta}\mathbf{y}_{t-1}^t\right) = 0$ (which is one of the FOCs), and so by stationarity $E\left(y_{T+1}^{T+2} - \beta_0^* - \boldsymbol{\beta}^*\mathbf{y}_T^{T+1}\right) = 0$. The unbiasedness of the forecasts holds irrespective of whether or not the revisions are zero mean, provided that $f = 1$ when revisions are non-zero mean. Suppose the goal were instead to forecast y_{T+1}^{T+1+f}, when $f > 1$, then the estimation and out-of-sample loss criteria no longer match, and RTV estimation would yield a systematic forecast error if $E\left(y_{T+1}^{T+1+f} - y_{T+1}^{T+2}\right) \neq 0$.[8] However, a simple solution suggests itself - the forecast should be corrected by an estimate of the difference between the vintage we wish to forecast and the first release, e.g., the sample mean of $y_t^{t+f} - y_t^{t+1}$, $t = 1, \ldots, T+1-f$.

Summarizing, the use of RTV data to estimate and forecast with AR(p) models in real time delivers forecasts that minimise the real-time expected loss (optimal forecasts) and unbiased forecasts of y_{T+1}^{T+2}. Unbiasedness also holds for $f > 1$ if data revisions are zero mean, but in the event of non-zero mean revisions a simple correction can be applied to the intercept.

6 Measuring the impact of using EOS and RTV data

In section 4, we showed that the use of EOS data to compute estimates of the AR model does not generate optimal forecasts. The differences between the optimal parameter values $(\phi_0^*, \boldsymbol{\phi}^*)$ in (8) and (13) and those obtained using EOS data ($(\alpha_0^*, \boldsymbol{\alpha}^*)$ in (11) and (14)) depend on the sizes of the means and variances of the various vintages of revisions, as well as on the properties of the data generating process for the underlying true data (e.g., $\mu_{\tilde{y}}$ and $\sigma_{\tilde{y}}^2$, and the correlation structure). In this section we firstly evaluate the analytical formulae numerically for empirically relevant values of the key parameters, to assess the potential importance of using RTV versus EOS data for forecasting with AR models. Secondly, because the results derived in sections 4 and 5 were based on large sample approximations, we also assess by Monte Carlo the relevance of the analytical predictions for the 'small samples' that are likely to be used in practice.

[8] The expected forecast error is $E\left(y_{T+1}^{T+f+1} - y_{T+1}^{T+2}\right) = E\left(v_{T+1}^{T+f+1} - v_{T+1}^{T+2} + \varepsilon_{T+1}^{T+f+1} - \varepsilon_{T+1}^{T+2}\right) = -\sum_{i=f}^{l} \mu_{v_i} + \sum_{i=1}^{l} \mu_{v_i} + \left(\mu_{\varepsilon_1} - \mu_{\varepsilon_f}\right) = -\sum_{i=1}^{f-1} \mu_{v_i} + \left(\mu_{\varepsilon_1} - \mu_{\varepsilon_f}\right)$, where $E\left(\varepsilon_t^{t+i}\right) = -\mu_{\varepsilon_i}$, and $E\left(v_t^{t+1}\right) = -\sum_{i=1}^{l} \mu_{v_i}$ and $E\left(v_t^{t+l}\right) = -\mu_{v_l}$, and so would be non-zero if either $\mu_{\varepsilon_f} \neq \mu_{\varepsilon_1}$ under noise revisions, or if $\sum_{i=1}^{f-1} \mu_{v_i} \neq 0$ when revisions are news.

6.1 Calibrating the key parameters

Because we want to assess the potential size of the forecast losses from using EOS data for typical macroeconomic aggregates subject to revisions, Figure 1 displays some of the key characteristics of the revisions processes for real output growth and two measures of inflation (GDP deflator, PCE deflator).[9] All three series are expressed in quarterly percentage differences. In the first (top left) panel, we plot the mean of each revision as a proportion of the mean of the first-released data. The figure plots the sample averages of revisions defined as $r_t^{(i)} = y_t^{t+1+i} - y_t^{t+i}$, for $i = 1, ..., 14$, calculated for the full period (1965Q3 onwards)[10]. It is apparent that the first revision tends to increase the mean of the first-released data by around 6% for output growth and 3% for GDP inflation. After the first revision, subsequent revisions tend to be smaller in terms of first-moment effects. In the second (top right) panel, we present the standard deviation of each revision as a proportion of the standard deviation of the first released data. For output growth and GDP inflation, the first-revision has a variance that is almost 25% of the first-release data. A salient feature is that the standard error of revisions tends to decrease in i. But it is important to recall that beyond the first revision we are calculating means and standard deviations by averaging over some revisions which are zero, because of the seasonality of revisions. This exacerbates the reduction in the variance of the second revision relative to that of the first revision.

The bottom left panel is the same as the top right, only restricted to the sub-period of the Great Moderation (assumed to be 1985 onwards). It is evident that the relative importance of revisions increases after 1985, as the lower variability of the measured first-release series is not matched one-for-one with a reduction in the variability of revisions. The first-revision standard deviation is now 40% of the standard deviation of the first-release data.

The bottom right panel is as for the top right, but based only on data for which t falls in the second quarter (i.e., the first release is in the 3rd quarter), and illustrates the seasonal pattern

[9] All real-time data employed in this paper are from the Real Time Data Set of the Philadelphia Fed available at the Philadelphia Fed webpage,
http://www.phil.frb.org/research-and-data/real-time-center/real-time-data/. See Croushore and Stark (2001).

[10] We plot $r_t^{(i)}$ for $i = 1$ to 14, as $r_t^{(14)}$ will account for all the 'regular' revisions.

(e.g., the second revision is zero). A similar figure would result if we chose any of the other three quarters. The smoothness in the top right panel comes from averaging over revisions that relate to observations that fall in different quarters of the year.

Based on these estimates of the typical sizes of revisions, we construct eight sets of parameters for the statistical model in section 3, assuming throughout that $p = 2$ and $l = 14$. Firstly, consider the AR coefficients of the model for $\widetilde{y}_t$. The first four sets of parameters in Table 1 have autoregressive coefficients that sum up to 0.4, while those of the last four sets of parameters sum to 0.8. The first block is more typical of a process such as output growth which exhibits moderate persistence, while the second is typical of a more persistent process such as inflation. Comparisons between these two blocks will therefore be informative about whether the relative forecast accuracy of EOS and RTV estimation depends upon the persistence of the underlying process.

For all sets of parameters, we assume that the means of the first and the fifth revisions are non-zero, as suggested by Figure 1. The mean values were chosen in conjunction with the values of the AR coefficients to give ratios of revision means to first-released data of 4% and 2%, as suggested by Figure 1. Within each of the two blocks we also vary the sizes of the standard deviations of the revisions (σ_{r_i}) compared to the standard deviation of the first-release data ($\sigma_{y_t^{t+1}}$). The first set of values within each block has a relatively large first revision variance ($\sigma_{r_1}/\sigma_{y_t^{t+1}} = .4$), followed by equal-sized revisions of smaller variance ($\sigma_{r_i}/\sigma_{y_t^{t+1}} = .2$, for $i = 2, ...13$), with a small final revision ($\sigma_{r_{14}}/\sigma_{y_t^{t+1}} = .1$). This decay is a stylized representation of the results in the second panel of Figure 1.[11] The second set of revision standard errors has proportionally larger revision standard deviations (50% larger) than the first set, and might be motivated by considering the Great Moderation period (as opposed to the full sample). The third one is useful to check the effect of no decay in the standard errors of the revisions. Finally, the fourth one assumes that the last revision reveals the true data ($\sigma_{r_{14}} = 0$).

The parameter values set out in Table 1 will be applied under the assumption that revisions are news, and under the assumption of noise. This will allow us to determine whether the news

[11] Specifically, for the purpose of computing the DGP parameters, we use $\sigma^2_{y_t^{t+1}} = \frac{R_1^2(1-\rho_2)}{(1+\rho_2)[(1-\rho_2)^2-\rho_1^2]}$.

versus noise issue is relevant to the relative forecast accuracy of EOS and RTV estimation of AR models for forecasting, both in large samples and in small samples when we allow for parameter estimation uncertainty in the Monte Carlo simulations. Because we have carefully calibrated our design parameters to reproduce the sorts of patterns we observe in the revisions of output growth and inflation, we would hope that our results might be informative about empirical outcomes.

6.2 Numerical quantification of gains to RTV versus EOS estimation: large sample results

In this section we evaluate the impact of data revisions on estimating and forecasting with an AR model assuming a large sample, such that the asymptotic results hold. For each set of parameter values in Table 1, we allow for revisions to be either pure news ($\sigma_{r_i} = \sigma_{v_i}$) or pure noise ($\sigma_{r_i} = \sigma_{\varepsilon_i}$). We compute the population values (ϕ_0^*, ϕ_1^*, ϕ_2^*) and (α_0^*, α_1^*, α_2^*) using equations (8)-(13), and (11)-(14), respectively. We then compute the population value of the loss function, (7), with $f = 1$, at these parameter values, and also calculate the bias of EOS estimation (from Remark 2). The results are recorded in Table 2. We find that the impact of data revisions is larger for more persistent data (compare DGPs 5 to 8 versus 1 to 4), and that there are marked changes in individual autoregressive coefficients (e.g., under news, for DGP 7 $\phi_1^* = 0.69$ and $\alpha_1^* = 0.51$), but that the sum of the autoregressive parameters changes less. The size of the forecast bias under EOS estimation is commensurate with the sizes of the revision means, while the loss in terms of MSFE from EOS estimation in comparison with optimal forecasting is in the 2-3% range for the more persistent data (DGPs 5 to 8) but it is smaller for the less persistent data.

In terms of the effect of the pattern of the variances of the revisions on the relative performance of EOS versus RTV estimation, we find that: i) under noise, EOS is more heavily penalized when the variance of the first revision is large relative to subsequent revisions (DGP 6); and ii) under news revisions, EOS fares relatively poorly in the same scenario, but is worst affected when there is no variance decay so that $\sigma_{v_i} = 0.3$ for $i = 1, \ldots, l$ (DGP 7). When the last vintage reveals the true data (DGPs 4 and 8), our main results still hold, so that even when EOS estimation is based on the true data (for all but the last 3-years of observations) one would still do better to use RTV estimation.

6.3 Monte Carlo estimates of small sample effects

In this section we assess by Monte Carlo simulation whether the large-sample results provide a useful guide to small-sample outcomes. Specifically, we assess whether: i) the values $(\phi_0^*,\ \phi_1^*, \phi_2^*)$ and $(\alpha_0^*,\ \alpha_1^*, \alpha_2^*)$ are good approximations to the estimates obtained using RTV and EOS data with small samples; and ii) whether greater or smaller losses to using EOS estimation relative to RTV estimation data are realised in small samples. We again use the set of parameters detailed in Table 1, and assuming that revisions are either news or noise, we simulate data using the statistical model described in section 3, and forecast from the AR model estimated by both EOS and RTV.

Table 3 presents the average bias across replications of estimating the autoregressive parameters using RTV $((\hat{\beta}_0,\ \hat{\beta}_1, \hat{\beta}_2)$ versus $(\phi_0^*,\ \phi_1^*, \phi_2^*))$ and EOS data $((\hat{\alpha}_0,\ \hat{\alpha}_1, \hat{\alpha}_2)$ versus $(\alpha_0^*,\ \alpha_1^*, \alpha_2^*))$. The entries are computed with samples of size $T = 50, 100, 200$ and 500, and employ $10,000$ replications. The asymptotic results are seen to provide a reasonable approximation to the small-sample estimates of the slope parameters with samples of size 200, but larger samples are required for the intercepts.

Table 4 presents the differences of the absolute value of the forecast biases between forecasts using EOS and RTV data for four different sample sizes. The first column records the differences in forecast biases computed in Table 2 for ease of comparison. Table 4 also shows the MSFE of forecasts computed with RTV data as a ratio of the MSFE of forecasts computed with EOS data. Finally, using the same estimates employed to compute one-step-ahead forecasts, we compute four-step-ahead forecasts by iteration. The right panel of Table 4 presents MSFE ratios of these forecasts, where the actuals are again the first-release values (that is, y_{T+4}^{T+1+4}). Our analytical formulae generally work well for the bias differences and MSFE ratios when the sample is large $(T = 500)$. An important finding for practical forecasting is that the losses to using EOS instead of RTV data are markedly larger for small samples when revisions are news, so that the analytical results downplay the likely empirical relevance of RTV estimation. For example, consider the relative performance of EOS and RTV estimation in terms of MSFE for DGP 6 when there are news revisions. For $T = 50$ the gains to RTV are around 7%, while the large sample results indicate gains closer to 4%. The results for forecasting four-steps ahead suggest gains of 3 to 4% for the more persistent processes (DGPs 5 to 8) when there are news revisions, for $T = 50$.

In summary, we would expect that the forecast loss of using EOS data instead of RTV data would be higher when i) the estimation sample is relatively small, ii) the process is reasonably persistent, and iii) revisions primarily add news.

6.4 Monte Carlo results based on a VAR of multiple vintages

In order to see whether our findings are robust to the way we have modelled the data revisions process, we also experiment with a version of the vintage-balanced VAR (VB-VAR) of Hecq and Jacobs (2009). The VB-VAR is closely related to the VAR models used by Garratt *et al.* (2008, 2009)) to model real-time data. The VB-VAR is given by:

$$\mathbf{z}_t = \mathbf{c}_0 + \sum_{i=1}^{p} \mathbf{\Gamma}_i \mathbf{z}_{t-i} + \boldsymbol{\varepsilon}_t \tag{16}$$

where $\mathbf{z}_{t-i} = \left[y_{t-1-i}^{t-i}, y_{t-2-i}^{t-i}, \ldots, y_{t-q-i}^{t-i} \right]'$, $i = 0, 1, \ldots, p$, $\mathbf{c}_0$ is $q \times 1$, $\mathbf{\Gamma}_i$ is $q \times 1$, and $\boldsymbol{\varepsilon}_t$ is $q \times 1$. The vector $\mathbf{z}_t$ (with $q = l$) differs from the vector $\mathbf{y}_t$ from the statistical framework described in section 3, because the data are organised by balancing the panel of data vintages as suggested by Hecq and Jacobs (2009). This way of employing the panel of data vintages means that the forecasts will be conditioned on the latest available estimates of the current and lagged observations. For example, for $p = 1$ the forecast of y_{T+1}^{T+1+f} will be the first element of the vector $\hat{\mathbf{z}}_{T+2} = \widehat{\mathbf{c}}_0 + \widehat{\mathbf{\Gamma}}_1 \mathbf{z}_{T+1}$. The VAR model captures the dynamics between different vintage estimates of the same observations, but does not clearly disentangle news and noise revisions.

The third panel of Table 4 reports the relative performance of AR models estimated with RTV and EOS data when data are generated from an estimated VB-VAR. The model was estimated on output growth, GDP deflator inflation, and PCE deflator inflation, giving rise to the three sets of results reported in the table (rows 9, 10 and 11). In each case, the VB-VAR was specified with $p = 1$, $q = 14$. We find gains to RTV over EOS of the order of 2 to 4% for $T = 50$, depending on the DGP, rising to 7% when $T = 500$ for the PCE deflator VB-VAR DGP.

The VB-VAR model can also be used to assess whether AR models are competitive with models exploiting multiple vintages in terms of real-time forecasting. The VB-VAR model requires the estimation of $(q^2 + q)$ parameters, which might be expected to adversely affect its forecast performance in small samples. The last panel of Table 4 presents a comparison of the forecast accuracy of the VB-VAR model (with $q = 14$) against the AR with EOS data, when the data generating process is the VB-VAR model calibrated to a US macro variable. Gains to using multiple vintages when the sample is large ($T = 500$) are around 11-15%, but the losses are three times larger when the sample is short ($T = 50$). These results suggest that parameter uncertainty may seriously curtail the value of multiple-vintage models in short samples.

7 Forecasting US output growth and inflation

Our first empirical forecasting exercise compares the forecast performance of the use of RTV and EOS data for simple autoregressive models for predicting quarterly output growth and inflation 1 and 4-steps ahead. In addition, we include the VB-VAR as the multiple-vintage model to see whether the use of multiple data vintages improves forecast accuracy. We begin with the autoregressive models (rather than ADL models) because the analysis in sections 4 and 5 that relates forecast performance to the properties of revisions is more directly applicable.

We then present a forecasting exercise for output growth and inflation that allows for explanatory variables in addition to the autoregressive terms. This parallels the pseudo out-of-sample forecasting exercises of Stock and Watson (2003, 2008) for these two variables, albeit using only a handful of candidate explanatory variables. Nevertheless, we are able to assess whether the theoretical advantages to the use of RTV over EOS in real-time forecasting are realized in practice. We will also consider very short horizon forecasts (or 'nowcasts') as Koenig *et al.* (2003) found marked gains to RTV-estimation of models for quarterly output growth using monthly indicators at such horizons.

7.1 AR models of output growth and inflation: RTV, EOS and the VB-VAR.

In this subsection we assess the empirical relevance of the analytical and Monte Carlo results for forecasts from autoregressive models of output growth and the two measures of inflation. The variables are defined as (one hundred times) the quarterly difference of the log of the level. We compute descriptive statistics of the different vintages of data and the revisions between them as follows. We consider first-released data y_t^{t+1}, data available three and a half years later y_t^{t+14}, as well as latest-available, which in our case is from the 2009:Q1 vintage dataset, denoted $y_t^{09:1}$. Table 5 presents means, standard deviations and first-order autocorrelations for the three data series, as well as p-values of tests for whether revisions ($y_t^{t+14} - y_t^{t+1}$ and $y_t^{09:1} - y_t^{t+1}$) are noise, or add news, and whether they are zero-mean. These are all calculated for the forecast period 1985:Q3 to 2006Q4, and separately for the estimation period 1965:Q3 to 1985:Q2.[12] Recall that revisions are defined as noise if the initial estimate is an observation on the final series but measured with error, so that the revisions are uncorrelated with final value, but are correlated with data available when the initial estimate was made. Hence noisy revisions are predictable. Alternatively, revisions are news if the initial estimate is an efficient forecast of the final value, such that the revision is unpredictable from information available at the time the initial estimate was made. We test for news and noise revisions using, respectively, the following auxiliary regressions:

$$y_t^{t+l} - y_t^{t+1} = \alpha + \beta y_t^{t+1} + \omega_t$$

$$y_t^{t+l} - y_t^{t+1} = \alpha + \beta y_t^{t+l} + \omega_t$$

[12] Real-time data are available starting with the 1965:Q4 vintage.

where the null hypothesis is that $\alpha = \beta = 0$ in both cases. In place of y_t^{t+l}, we use both y_t^{t+14} and y_t^{09Q1}. We also tests separately whether revisions are zero mean (H_0: $\alpha = 0$ in $y_t^{t+l} - y_t^{t+1} = \alpha + \omega_t$).

Data revisions to output growth and inflation are seen to have different characteristics, and show some variation across the forecast and estimation periods. For output growth we can reject the noise hypothesis for both periods using the 2009:1 data vintage, and there is no evidence against the news hypothesis using the $t + 14$ data vintage (the latter matching the findings of Mankiw and Shapiro (1986)). For both inflation measures the revisions relative to the 2009:1 vintage can be assumed to be noise over the forecast period, and news over the estimation period, while the revisions relative to y_t^{t+14} cannot be so easily categorised. There is also evidence that the 2009:1 revisions to output growth have been significantly upward, as have the $t + 14$ revisions to both inflation rates over the forecast period.

The results of sections 4 and 5 suggest that we should expect improvements in forecast accuracy from using RTV instead of EOS data when there are news revisions, especially when estimation sample sizes are short, as well as when there are noise revisions for more highly persistent processes (as is true of inflation compared to output growth). Our out-of-sample period for forecast comparison is such that the initial estimation sample has around 100 observations, and as we adopt a recursive forecasting scheme, this increases to around 194 by the end.

We compare forecasts computed with RTV and EOS data for the three series, using autoregressive models with fixed specifications - the lag length for output growth is 1, and for the inflation series, 4. These autoregressive orders were chosen based on the performance of these models when forecasting with final data. Recall that both RTV and EOS condition the forecasts on exactly the same information set $(y_T^{T+1}, \ldots, y_{T-p+1}^{T+1})$ for T =1985:Q3,...,2008:Q4. The forecasts will differ to the extent that the parameters of the AR models estimated with EOS and RTV data differ. As the forecast origin is moved through the data, the AR model using EOS data will be estimated using the full set of data from the new vintage that becomes available at that origin. By way of contrast, with RTV data, we add only the estimate of the most recent period from the new vintage, as described in section 5.

We evaluate forecasts by computing forecast errors using first-released data (y_{T+1}^{T+2}), data after three and a half years of revision (y_{T+1}^{T+15}) and data from the last available vintage (y_{T+1}^{09Q1}). Table 6 presents ratios of RMSFEs of forecasting using RTV and EOS data such that values smaller than one favour RTV. We give results for one-step-ahead forecasts, and results for four-step-ahead forecasts (computed by iteration of the one-step forecasts). In general there are gains to RTV over EOS for all three series, and some of these are of the order of 4%. There are gains to the use of RTV data for forecasting the early release as well as the revised data, and they are generally larger for four-step-ahead forecasts. Bias-correction of the RTV forecasts (as discussed in section 5) is not effective, as might have been anticipated from the differences in the mean revisions between the estimation and forecast periods. Nevertheless, there is promising evidence in favour of using RTV data instead of EOS data.

Can we do better still by making use of multiple vintage estimates of the same observations, as in the VB-VAR model? Table 6 also reports the ratios of the RMSFEs of the VB-VAR model forecasts to the AR (estimated with EOS data), for a lag order of 1, but experimenting with different values of q (the number of vintages). In general, gains from using VB-VAR with respect to the AR estimated with EOS data are similar in size to the gains from using RTV data for AR forecasting, and in many cases the VB-VAR performance is worse. An exception is forecasting the GDP deflator inflation: RMSFEs reductions of 5% are observed when predicting y_{T+1}^{T+15}, especially when q is large. In short, using multiple data vintages does not result in clearly superior forecasts across the three variables taken together. In this regard the empirical forecasting results are less supportive of the VB-VAR than the Monte Carlo reported in table 4. This likely reflects the smaller sample sizes that are available empirically, and non-constancies in the correlation structure of the different vintage estimates over time (assumed absent in the Monte Carlo by construction).

7.2 ADL models of output growth and inflation

In this section we compare the relative forecasting ability of autoregressive distributed lag models estimated by EOS and RTV. For output growth, we consider two indicators: industrial production and employment. For the inflation variables, the indicators are the same two variables with the addition of output growth. The choice of these indicators is supported by their popularity in forecasting exercises that aim at assessing the predictive power of economic activity variables for output and inflation and also by the availability of real-time data in the Philadelphia dataset.[13]

We consider forecasts of quarterly growth (at an annual rate) 1 and 4 quarters ahead, as before, and in addition we generate 'nowcasts': in this case the horizon is $h = 0$. The ADL models specified for nowcasting correspond closely to the distributed lag models of Koenig *et al.* (2003) (who regressed quarterly output growth on the contemporaneous and 4 lags of the indicator sampled monthly, but preliminary investigation suggested nothing is lost by using the contemporaneous and one lag of the indicator sampled quarterly instead). We also consider two in-sample periods. The first begins in 1959Q1, as in the previous sub-section; the second is the much shorter estimation period beginning in 1979Q1, to be comparable with Koenig *et al.* (2003).

The implementation of the RTV approach for ADL(p_y, p_x) models at each forecast horizon h ($h \geq 1$) follows:

$$y_t^{t+1} = \beta_0 + \sum_{i=0}^{p_y-1} \beta_{(1+i)} y_{t-h-i}^{t+1-h} + \sum_{i=0}^{p_x-1} \gamma_i x_{t-h-i}^{t+1-h} + e_{RTV,t} \tag{17}$$

for $t = \max(p_y + h, p_x + h) + 1, ..., T$, where p_y is the autoregressive order and p_x is the number of lags of the indicator. In the case of nowcasting ($h = 0$), we use:

$$y_t^{t+1} = \beta_0 + \sum_{i=1}^{p_y} \beta_i y_{t-i}^{t} + \sum_{i=0}^{p_x-1} \gamma_i x_{t-i}^{t+1} + e_{RTV,t}.$$

Consider firstly the results for forecasting output growth reported in table 7. The entries which report the relative RMSFEs of the models estimated with RTV and EOS data suggest gains of around 5% at the $h = 0$ horizon for both industrial production and employment as the explanatory variable *provided* we restrict the start of the estimation sample to 1979Q1, as in Koenig *et al.* (2003). On the longer estimation sample RTV is no better that EOS. This primarily reflects a worsening in

[13] Quarterly vintages of industrial production and employment are constructed by taking the first month of the quarter vintages. We obtain quarterly series from the monthly by averaging over the months in the quarter.

the EOS-based forecasts when the sample is shortened, with the RTV-generated forecasts being less affected by sample size. We also compute the ratio of the RMSFEs of the ADL model to the AR (both estimated by RTV), and only for the $h = 0$ horizon do we find that the indicator variables enhance forecast performance.

In terms of the two inflation series, we get a clear picture. The use of RTV data improves the forecast accuracy of the ADL models at all horizons (gains of around 5%) in comparison with EOS, albeit that the gains at $h = 4$ tend to depend on the shorter sample being used. But compared to the AR(4) benchmark, the ADL models do not generally improve the forecasts of inflation.

To conclude: the forecast accuracy of the ADL models is generally enhanced by RTV estimation, although for output growth this result is dependent on truncating the available estimation sample. Nevertheless, the AR benchmark models (estimated using RTV data) prove more competitive against the ADL models except for nowcasting output growth. The results of our real-time forecasting exercise are in tune with those of the pseudo out-of-sample exercises of Stock and Watson (2003, 2008) who report that indicator variables do not consistently enhance output growth and inflation forecasts over the period from 1985 onwards.

8 Conclusions

In recent times there has been a growing appreciation of the effects of data revisions on various aspects of macro-modelling (such as the calculation of output gaps and conduct of monetary policy: e.g., Orphanides (2001) and Orphanides and van Norden (2005)), as well as the relevance of data revisions for forecasting (as reviewed by Croushore (2006)). We have tackled one aspect of forecasting when there are data revisions: namely, can we improve real-time forecasts of autoregressive models by exploiting better the real-time data currently available? Our analytical results show that the use of the latest available vintage at each point in time (EOS) will lead to AR model forecasts which do not minimise the real-time forecasting loss function. Moreover, the use of EOS data may result in biased forecasts when data revisions are non-zero mean. The use of early-release data in estimation (RTV) improves forecast accuracy in population. When data revisions are non-zero mean, RTV-forecasts remain unbiased when the aim is to forecast the first-release data. If the goal is to forecast post-revision data, the logic of the RTV approach suggests a simple bias-correction can be applied to the forecasts.

When we use AR models for forecasting output growth and inflation, we observe the gains to using RTV data predicted by our analytical results and the Monte Carlo simulations. Forecasts from the AR model estimated with RTV data are competitive with the vintage-based VAR model (VB-VAR), suggesting that using more data vintages is not advantageous for forecasting post WWII US output growth and inflation, although other models might result in different conclusions. We also observe improvements in forecasting accuracy from using RTV data when estimating autoregressive distributed lag models for predicting output growth and inflation using economic indicators. However, our real-time forecasting exercise does not change the conclusions we draw regarding the predictability of inflation using Phillips curve-type models with 'activity variables' compared to the pseudo out–of-sample exercise of Stock and Watson (2008).

Finally, one might wonder whether alternatives to the VB-VAR model offer better ways of harnessing the information content in past data vintages for forecasting. It is not clear that more elaborate, complicated models that simultaneously model the true process and the revisions process would be expected to yield more accurate forecasts, given the findings of the recent empirical forecast comparison literature that 'simple' models generate competitive forecasts relative to more sophisticated models,[14] but this would be an interesting topic to explore.

References

Aruoba, S. B. (2008). Data revisions are not well-behaved. *Journal of Money, Credit and Banking*, **40**, 319–340.

Clements, M. P., and Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data:

[14]See, for example, Fildes and Makridakis (1995), Clements and Hendry (1999) and Makridakis and Hibon (2000).

Forecasting output growth in the United States. *Journal of Business and Economic Statistics*, **26**, 546–554. No. 4.

Clements, M. P., and Galvão, A. B. (2009). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics*, **24**, 1187–1206.

Clements, M. P., and Hendry, D. F. (1999). *Forecasting Non-stationary Economic Time Series.* Cambridge, Mass.: MIT Press.

Corradi, V., Fernandez, A., and Swanson, N. R. (2009). Information in the revision process of real-time datasets. *Journal of Business and Economic Statistics*, **27**, 455–467.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. In Elliott, G., Granger, C., and Timmermann, A. (eds.), *Handbook of Economic Forecasting, Volume 1. Handbook of Economics 24*, pp. 961–982: Elsevier, Horth-Holland.

Croushore, D., and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, **105**, 111–130.

Croushore, D., and Stark, T. (2003). A real-time data set for macroeconomists: Does the data vintage matter?. *The Review of Economics and Statistics*, **85**, 605–617.

Cunningham, A., Eklund, J., Jeffery, C., Kapetanios, G., and Labhard, V. (2009). A state space approach to extracting the signal from uncertain data. *Journal of Business & Economic Statistics.*

Diebold, F. X., and Rudebusch, G. D. (1991a). Forecasting output with the composite leading index: A real-time analysis. *Journal of the American Statistical Association*, **86**, 603–610.

Diebold, F. X., and Rudebusch, G. D. (1991b). Turning point prediction with the composite leading index: An ex ante analysis. In Lahiri, K., and Moore, G. H. (eds.), *Leading economic indicators: New approaches and forecasting records*, pp. 231–256: Cambridge University Press, Cambridge.

Faust, J., Rogers, J. H., and Wright, J. H. (2003). Exchange rate forecasting: The errors we've really made. *Journal of International Economic Review*, **60**, 35–39.

Fildes, R. A., and Makridakis, S. (1995). The impact of empirical accuracy studies on time series analysis and forecasting. *International Statistical Review*, **63**, 289–308.

Fixler, D. J., and Grimm, B. T. (2005). Reliability of the NIPA estimates of U.S. economic activity. *Survey of Current Business*, **85**, 9–19.

Fixler, D. J., and Grimm, B. T. (2008). The reliability of the GDP and GDI estimates. *Survey of Current Business*, **88**, 16–32.

Garratt, A., Lee, K., Mise, E., and Shields, K. (2008). Real time representations of the output gap. *Review of Economics and Statistics*, **90**, 792–804.

Garratt, A., Lee, K., Mise, E., and Shields, K. (2009). Real time representations of the UK output gap in the presence of model uncertainty. *International Journal of Forecasting*, **25**, 81–102.

Harrison, R., Kapetanios, G., and Yates, T. (2005). Forecasting with measurement errors in dynamic models. *International Journal of Forecasting*, **21**, 595–607.

Harvey, A. C., McKenzie, C. R., Blake, D. P. C., and Desai, M. J. (1983). Irregular data revisions. In Zellner, A. (ed.), *Applied Time Series Analysis of Economic Data*, pp. 329–347: US Department of Commerce, Washington D.C., Economic Research Report ER-5.

Hecq, A., and Jacobs, J. P. A. M. (2009). On the VAR-VECM representation of real time data. Discussion paper, mimeo, University of Maastricht, Department of Quantitative Economics.

Howrey, E. P. (1978). The use of preliminary data in economic forecasting. *The Review of Economics and Statistics*, **60**, 193–201.

Howrey, E. P. (1984). Data revisions, reconstruction and prediction: an application to inventory investment. *The Review of Economics and Statistics*, **66**, 386–393.

Jacobs, J. P. A. M., and van Norden, S. (2007). Modeling data revisions: Measurement error and dynamics of 'true' values. Technical report cref 07-09, HEC, Montreal.

Kishor, N. K., and Koenig, E. F. (2010). VAR estimation and forecasting when data are subject to revision. *Journal of Business and Economic Statistics*. Forthcoming.

Koenig, E. F., Dolmas, S., and Piger, J. (2003). The use and abuse of real-time data in economic forecasting. *The Review of Economics and Statistics*, **85(3)**, 618–628.

Landefeld, J. S., Seskin, E. P., and Fraumeni, B. M. (2008). Taking the pulse of the economy. *Journal of Economic Perspectives*, **22**, 193–216.

Makridakis, S., and Hibon, M. (2000). The M3 Competition: results, conclusions and implications. *International Journal of Forecasting*, **16**, 451–476.

Mankiw, N. G., and 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.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review*, **91(4)**, 964–985.

Orphanides, A., and van Norden, S. (2005). The reliability of inflation forecasts based on output gaps in real time. *Journal of Money, Credit and Banking*, **37**, 583–601.

Patterson, K. D. (1995). An integrated model of the data measurement and data generation processes with an application to consumers' expenditure. *Economic Journal*, **105**, 54–76.

Patterson, K. D. (2003). Exploiting information in vintages of time-series data. *International Journal of Forecasting*, **19**, 177–197.

Robertson, J. C., and Tallman, E. W. (1998). Data vintages and measuring forecast model performance. *Federal Reserve Bank of Atlanta Economic Review, Fourth Quarter*, 4–20.

Sargent, T. J. (1989). Two models of measurements and the investment accelerator. *Journal of Political Economy*, **97**, 251–287.

Siklos, P. L. (2008). What can we learn from comprehensive data revisions for forecasting inflation: Some US evidence. In Rapach, D. E., and Wohar, M. E. (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty. Frontiers of Economics and Globalization. Volume 3*, pp. 271–299: Emerald.

Stark, T., and Croushore, D. (2002). Forecasting with a real-time data set for macroeconomists.

Journal of Macroeconomics, **24**, 507–531.

Stock, J. H., and Watson, M. W. (2003). How Did Leading Indicator Forecasts Perform During the 2001 Recession. *Federal Reserve Bank of Richmond, Economic Quarterly*, **89/3**, 71–90.

Stock, J. H., and Watson, M. W. (2008). Phillips Curve Inflation Forecasts. Working paper 14322, NBER, Cambridge, MA.

A Proofs

Proof. Proposition 1: Optimal AR population parameter.

The expected squared forecast error is given by:

$$E\left(y_{T+1}^{T+1+f}-\phi'\mathbf{y}_T^{T+1}\right)^2=E\left(\rho_0+\boldsymbol{\rho}'\widetilde{\mathbf{y}}_T+R_1\eta_{1T+1}+\hat{v}_{T+1}^{T+1+f}+\varepsilon_{T+1}^{T+1+f}-\phi_0-\boldsymbol{\phi}'\widetilde{\mathbf{y}}_T-\boldsymbol{\phi}'\mathbf{v}_T^{T+1}-\boldsymbol{\phi}'\boldsymbol{\varepsilon}_T^{T+1}\right)^2$$

where $\mathbf{y}_T^{T+1}=\widetilde{\mathbf{y}}_T+\mathbf{v}_T^{T+1}+\boldsymbol{\varepsilon}_T^{T+1}$, $\hat{v}_{T+1}^{T+1+f}=-v_{T+1}^{T+2}+v_{T+1}^{T+1+f}$, with $\mathbf{y}_T^{T+1\prime}=\left(y_T^{T+1},\ldots,y_{T-p+1}^{T+1}\right)$, $\widetilde{\mathbf{y}}_T'=(\widetilde{y}_T,\ldots,\widetilde{y}_{T-p+1})$, $\mathbf{v}_T^{T+1\prime}=\left(v_T^{T+1},\ldots,v_{T-p+1}^{T+1}\right)$, with typical element $v_{T-j}^{T+1}=-\sum_{i=j+1}^{l}\sigma_{v_i}\eta_{2,T-j,i}$, for $j<l$, and $v_{T-j}^{T+1}=\sigma_{v_i}\eta_{2,T-j,l}$, for $j\geq l$, and $\boldsymbol{\varepsilon}_T^{T+1\prime}=\left(\varepsilon_T^{T+1},\ldots,\varepsilon_{T-p+1}^{T+1}\right)$. Note that $\mathbf{y}_{T+1}^{T+1+f}=\widetilde{\mathbf{y}}_{T+1}+\mathbf{v}_{T+1}^{T+1+f}+\boldsymbol{\varepsilon}_{T+1}^{T+1+f}=\rho_0+\boldsymbol{\rho}'\widetilde{\mathbf{y}}_T+R_1\eta_{1T+1}-\mathbf{v}_{T+1}^{T+2}+v_{T+1}^{T+1+f}+\varepsilon_{T+1}^{T+1+f}$, so that $\hat{v}_{T+1}^{T+1+f}=-v_{T+1}^{T+2}+v_{T+1}^{T+1+f}=0$ when $f=1$. We let $E\left(\widetilde{\mathbf{y}}_T\widetilde{\mathbf{y}}_T'\right)=\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}}+\boldsymbol{\mu}_{\widetilde{y}}\boldsymbol{\mu}_{\widetilde{y}}'$, where $\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}}=Var\left(\widetilde{\mathbf{y}}_T\right)$, $\boldsymbol{\mu}_{\widetilde{y}}=\mathbf{i}\mu_{\widetilde{y}}=E\left(\widetilde{\mathbf{y}}_T\right)$, $\mathbf{i}$ a p-dimensional vector of 1's; $\boldsymbol{\Sigma}_{\mathbf{v}}\equiv Var\left(\mathbf{v}_T^{T+1}\mathbf{v}_T^{T+1\prime}\right)=E\left(\mathbf{v}_T^{T+1}\mathbf{v}_T^{T+1\prime}\right)=diag\left(\sum_{i=1}^{l}\sigma_{v_i}^2,,\ldots,\sum_{i=p}^{l}\sigma_{v_i}^2\right)$ for $p\leq l$, with terms of $\sigma_{v_l}^2$ on the diagonal for $p>l$; $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\equiv Var\left(\boldsymbol{\varepsilon}_T^{T+1}\boldsymbol{\varepsilon}_T^{T+1\prime}\right)=E\left(\boldsymbol{\varepsilon}_T^{T+1}\boldsymbol{\varepsilon}_T^{T+1\prime}\right)=diag\left(\sigma_{\varepsilon_1}^2,\sigma_{\varepsilon_2}^2,\ldots\sigma_{\varepsilon_p}^2\right)$, with $\sigma_{\varepsilon_s}^2=\sigma_{\varepsilon_l}^2$ for $s>l$, and $\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}\mathbf{v}}\equiv Var\left(\widetilde{\mathbf{y}}_T\mathbf{v}_T^{T+1\prime}\right)=E\left(\widetilde{\mathbf{y}}_T\mathbf{v}_T^{T+1\prime}\right)$. Solving the first-order conditions $\partial E\left(y_{T+1}^{T+f}-\boldsymbol{\phi}'\mathbf{y}_T^{T+1}\right)^2/\partial\boldsymbol{\phi}=\mathbf{0}$ and $\partial E\left(y_{T+1}^{T+f}-\boldsymbol{\phi}'\mathbf{y}_T^{T+1}\right)^2/\partial\phi_0=\mathbf{0}$ gives:

$$\begin{aligned}\boldsymbol{\phi}^* &= \left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}}+\Sigma_{\mathbf{v}}+\Sigma_{\widetilde{\mathbf{y}}\mathbf{v}}+\Sigma_{\widetilde{\mathbf{y}}\mathbf{v}}'+\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right)^{-1}\left(\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}}+\boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}\mathbf{v}}'\right)\boldsymbol{\rho}\\ \phi_0^* &= \left(1-\boldsymbol{\phi}^{*\prime}\mathbf{i}\right)\mu_{\widetilde{y}}.\end{aligned}$$

■

Proof. Proposition 2: Estimation of AR(p) with EOS data (zero-mean data revisions).

1. If we allow $T\to\infty$ for a fixed l, then the OLS estimator using EOS data is asymptotically equivalent to the estimator obtained from a regression using only fully-revised data. Denote this data by $\{y_t,y_{t-1}\ldots y_{t-p}\}$, where $y_t=\widetilde{y}_t+v_t^{t+l}+\varepsilon_t^{t+l}$, and $y_t=\widetilde{y}_t$ if the true data are eventually revealed. As the data $\left\{y_t,\mathbf{y}_{t-1}=(y_{t-1}\ldots y_{t-p})'\right\}$ are covariance stationary we can use the population moments (assuming that $T\to\infty$) to compute the OLS estimators as values that satisfy

$$(\alpha_0^*,\boldsymbol{\alpha}^*)=\underset{\alpha_0,\boldsymbol{\alpha}}{\arg\min}E\left(y_t-\alpha_0-\boldsymbol{\alpha}\mathbf{y}_{t-1}\right)^2,$$

which are:

$$\alpha_0^*=E\left(y_t\right)-\boldsymbol{\alpha}E\left(\mathbf{y}_{t-1}\right) \tag{18}$$

and

$$E\left(\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\right)\boldsymbol{\alpha}^*+\alpha_0^*E\left(\mathbf{y}_{t-1}\right)-E\left(y_t\mathbf{y}_{t-1}\right)=0. \tag{19}$$

Combining (18) and (19) gives the standard FOC for $\boldsymbol{\alpha}^*$,

$$Cov\left(\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\right)\boldsymbol{\alpha}=Cov\left(y_t\mathbf{y}_{t-1}\right) \tag{20}$$

where $Cov\left(\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\right) = E\left(\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\right) - E\left(\mathbf{y}_{t-1}\right)E\left(\mathbf{y}_{t-1}'\right)$, $Cov\left(y_t\mathbf{y}_{t-1}\right) = E\left(y_t\mathbf{y}_{t-1}\right) - E\left(y_t\right)E\left(\mathbf{y}_{t-1}\right)$.

2. The moments in (18) are obtained from $E\left(y_t\right) = E\left(\widetilde{y}_t + v_t^{t+l} + \varepsilon_t^{t+l}\right) = \mu_{\widetilde{y}}$, and $E\left(\mathbf{y}_{t-1}\right) = \mathbf{i}\mu_{\widetilde{y}}$, since $\mathbf{y}_{t-1} = \widetilde{\mathbf{y}}_{t-1} + \mathbf{v}_{t-1}^{t+l} + \boldsymbol{\varepsilon}_{t-1}^{t+l}$, with $\widetilde{\mathbf{y}}_{t-1} = [\widetilde{y}_{t-1}, \ldots, \widetilde{y}_{t-p}]'$, $\mathbf{v}_{t-1}^{t+l} = \left[v_{t-1}^{t+l}, \ldots, v_{t-p}^{t+l}\right]'$, $\boldsymbol{\varepsilon}_{t-1}^{t+l} = \left[\varepsilon_{t-1}^{t+l}, \ldots, \varepsilon_{t-p}^{t+l}\right]'$, so that:

$$\alpha_0^* = \left(1 - \boldsymbol{\alpha}^{*\prime}\mathbf{i}\right)\mu_{\widetilde{y}}. \tag{21}$$

For (20) we obtain:

$$Cov\left(\mathbf{y}_{t-1}\mathbf{y}_{t-1}'\right) = \Sigma_{\widetilde{y}} + \boldsymbol{\Sigma}_{\underline{\mathbf{v}}} + \boldsymbol{\Sigma}_{\underline{\boldsymbol{\varepsilon}}} + \Sigma_{\widetilde{y}v_l} + \Sigma_{\widetilde{y}v_l}' \tag{22}$$

where $\Sigma_{\widetilde{y}v_l} \equiv E\left[\left(\widetilde{\mathbf{y}}_{t-1} - E\left(\widetilde{\mathbf{y}}_{t-1}\right)\right)\left(\mathbf{v}_{t-1}^{t+l} - E\left(\mathbf{v}_{t-1}^{t+l}\right)\right)'\right]$, $\boldsymbol{\Sigma}_{\underline{\mathbf{v}}} = E\left[\mathbf{v}_{t-1}^{t+l} - E\left(\mathbf{v}_{t-1}^{t+l}\right)\left(\mathbf{v}_{t-1}^{t+l} - E\left(\mathbf{v}_{t-1}^{t+l}\right)\right)'\right]$ $\sigma_{v_l}^2 I_p$, $\boldsymbol{\Sigma}_{\underline{\boldsymbol{\varepsilon}}} = E\left[\boldsymbol{\varepsilon}_{t-1}^{t+l} - E\left(\boldsymbol{\varepsilon}_{t-1}^{t+l}\right)\left(\boldsymbol{\varepsilon}_{t-1}^{t+l} - E\left(\boldsymbol{\varepsilon}_{t-1}^{t+l}\right)\right)'\right] = \sigma_{\varepsilon_l}^2 I_p$. Note that $\Sigma_{\widetilde{y}v_l}$ is upper diagonal, and its diagonal is (minus) the diagonal of $\boldsymbol{\Sigma}_{\underline{\mathbf{v}}}$. Also:

$$Cov\left(y_t\mathbf{y}_{t-1}\right) = V\left(\widetilde{\mathbf{y}}_{t-1}\widetilde{\mathbf{y}}_{t-1}'\right)\rho + V\left(\mathbf{v}_{t-1}^{t+l}\widetilde{\mathbf{y}}_{t-1}'\right)\rho = \Sigma_{\widetilde{y}}\rho + \Sigma_{\widetilde{y}v_l}'\rho \tag{23}$$

Substituting (22) and (23) into (20) gives:

$$\boldsymbol{\alpha}^* = \left(\Sigma_{\widetilde{y}} + \boldsymbol{\Sigma}_{\underline{\mathbf{v}}} + \boldsymbol{\Sigma}_{\underline{\boldsymbol{\varepsilon}}} + \Sigma_{\widetilde{y}v_l} + \Sigma_{\widetilde{y}v_l}'\right)^{-1}\left(\Sigma_{\widetilde{y}} + \Sigma_{\widetilde{y}v_l}'\right)\rho. \tag{24}$$

Recall that $\boldsymbol{\Sigma}_{\underline{\boldsymbol{\varepsilon}}} \neq \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} = diag\left\{\sigma_{\varepsilon_1}^2, \ldots \sigma_{\varepsilon_p}^2\right\}$, $\Sigma_{\widetilde{y}v_l} \neq \boldsymbol{\Sigma}_{\widetilde{\mathbf{y}}\mathbf{v}}$ and $\boldsymbol{\Sigma}_{\underline{\mathbf{v}}} \neq \boldsymbol{\Sigma}_{\mathbf{v}}$ from the assumption that in the large samples the use of EOS data implies the use of data from the $t+l$ vintage. Hence when $\sigma_{v_l} = \sigma_{\varepsilon_l} = 0$, the use of EOS gives the same large-sample estimates of the model parameters as using the true data $\{\widetilde{y}_t\}$.

■

Proof. Remark 1: Forecasts obtained using EOS data to estimate the AR(p) model are unbiased when data revisions follow (2)-(4).

By construction, $\{\alpha_0^*, \boldsymbol{\alpha}^*\}$ satisfy $E\left(y_{T+1} - \alpha_0 - \boldsymbol{\alpha}'\mathbf{y}_T\right) = 0$, where y_{T+1} and $\mathbf{y}_T = \left(y_T, \ldots, y_{T-p+1}\right)'$ are 'final data'. But the forecasting exercise is to forecast y_{T+1} using the latest available data, $\mathbf{y}_T^{T+1} = \left(y_T^{T+1}, \ldots, y_{T-p+1}^{T+1}\right)'$. Consider the expected value of the forecast error $E\left(y_{T+1}^{T+2} - \alpha_0 - \boldsymbol{\alpha}'\mathbf{y}_T^{T+1}\right)$, where we substitute $y_{T+1}^{T+2} = y_{T+1} - \left(y_{T+1} - y_{T+1}^{T+2}\right)$ and $\mathbf{y}_T^{T+1} = \mathbf{y}_T - \left(\mathbf{y}_T - \mathbf{y}_T^{T+1}\right)$ to give:

$$E\left(y_{T+1}^{T+2} - \alpha_0 - \boldsymbol{\alpha}\mathbf{y}_T^{T+1}\right) = E\left(y_{T+1}^{T+2} - y_{T+1} - \boldsymbol{\alpha}'\left(\mathbf{y}_T^{T+1} - \mathbf{y}_T\right)\right) = 0$$

because $E\left(y_{T+1}^{T+2}\right) = E\left(y_{T+1}\right) = \mu_{\widetilde{y}}$ and $E\left(\mathbf{y}_T^{T+1}\right) = E\left(\mathbf{y}_T\right) = \mathbf{i}\mu_{\widetilde{y}}$. ■

Proof. Proposition 3: Optimal AR population parameter when data revisions are non-zero mean.

The optimal parameters solve $E\left(y_{T+1}^{T+2}\right) = \phi_0 + E\left(\boldsymbol{\phi}'\mathbf{y}_T^{T+1}\right)$. The optimal slope parameters $\boldsymbol{\phi}^*$ only depend on (centred) second moment matrices, so are the same whether the data revision process is (2)-(5)-(6) or (2)-(4). Under the data revision process (2)-(5)-(6), we have

$$E\left(y_{T+1}^{T+2}\right) = E\left(\widetilde{y}_t + v_t^{t+1} + \varepsilon_t^{t+1}\right) = \mu_{\widetilde{y}} - \sum_{i=1}^{l} \mu_{v_i} - \mu_{\varepsilon_1},$$

where $E\left(\boldsymbol{\phi}'\mathbf{y}_{t-1}^t\right) = \boldsymbol{\phi}'\left(\mathbf{i}\mu_{\widetilde{y}} + \boldsymbol{\mu}_\varepsilon + \boldsymbol{\mu}_v\right)$, with $\mathbf{i}\mu_{\widetilde{y}} = E\left([\widetilde{\mathbf{y}}_{t-1}]\right) = E\left([\widetilde{y}_{t-1}, \ldots, \widetilde{y}_{t-p}]'\right)$, $\mathbf{i}$ a p-dimensional vector of 1's, and $\boldsymbol{\mu}_\varepsilon$ and $\boldsymbol{\mu}_v$ are the means of the noise and news revisions, $\boldsymbol{\mu}_\varepsilon \equiv E\left([\varepsilon_{t-1}^t, \ldots, \varepsilon_{t-p}^t]\right) = \left[-\mu_{\varepsilon_1}, ..., -\mu_{\varepsilon_p}\right]$, $\boldsymbol{\mu}_v \equiv E\left([v_{t-1}^t, \ldots, v_{t-p}^t]\right) = \left[-\sum_{i=1}^{l} \mu_{v_i}, ..., -\sum_{i=p}^{l} \mu_{v_i}\right]$. Finally, $\mu_{\widetilde{y}} = (1-\rho(1))^{-1}\left[\rho_0 + \sum_{i=1}^{l} \mu_{v_i}\right]$. Using the previous results, it is straightforward to show that:

$$\phi_0^* = \left(1 - \boldsymbol{\phi}^{*\prime}\mathbf{i}\right)\mu_{\widetilde{y}} - \sum_{i=1}^{l} \mu_{v_i} - \mu_{\varepsilon_1} - \boldsymbol{\phi}^{*\prime}\boldsymbol{\mu}_\varepsilon - \boldsymbol{\phi}^{*\prime}\boldsymbol{\mu}_v.$$

■

Proof. Proposition 4: Estimation of AR(p) with EOS data (zero-mean data revisions).

Allowing for non-zero mean revisions, in place of $E(y_t) = \mu_{\widetilde{y}}$ and $E(\mathbf{y}_{t-1}) = \mathbf{i}\mu_{\widetilde{y}}$ in Proposition 2, we now have $E(y_t) = E\left(\widetilde{y}_t + v_t^{t+l} + \varepsilon_t^{t+l}\right) = \mu_{\widetilde{y}} - \mu_{v_l} - \mu_{\varepsilon_l}$, where $E(\mathbf{y}_{t-1}) = \mathbf{i}\mu_{\widetilde{y}} - \mathbf{i}(\sum_{i=1}^{l} \mu_{v_i}) - \mathbf{i}\,\mu_{\varepsilon_1}$. $\boldsymbol{\alpha}^*$ is still given by (24), but (21) becomes:

$$\alpha_0^* = \left(1 - \boldsymbol{\alpha}^{*\prime}\mathbf{i}\right)\left(\mu_{\widetilde{y}} - \mu_{v_l} - \mu_{\varepsilon_l}\right).$$

■

Proof. Remark 2: Forecasts computed using EOS data to estimate AR(p) model are biased when data revisions are described by (2)-(5)-(6).

Consider the first moment properties of the forecast errors when an AR(p) is estimated by EOS in the presence of non-zero mean data revisions. Suppose the aim is to forecast the first vintage estimate y_{T+1}^{T+2}. To see that EOS population forecasts are generally biased, note that by construction, $\{\alpha_0^*, \boldsymbol{\alpha}^*\}$ satisfy $E(y_{T+1} - \alpha_0 - \boldsymbol{\alpha}'\mathbf{y}_T) = 0$, where y_{T+1} and $\mathbf{y}_T = (y_T, \ldots, y_{T-p+1})$ are 'final data'. But forecasts are of necessity conditioned on $\mathbf{y}_T^{T+1} = \left(y_T^{T+1}, \ldots, y_{T-p+1}^{T+1}\right)$, so of interest is the expected value of the forecast error $E\left(y_{T+1}^{T+2} - \alpha_0^* - \boldsymbol{\alpha}^{*\prime}\mathbf{y}_T^{T+1}\right)$. Substituting $y_{T+1}^{T+2} = y_{T+1} - \left(y_{T+1} - y_{T+1}^{T+2}\right)$ and $\mathbf{y}_T^{T+1} = \mathbf{y}_T - \left(\mathbf{y}_T - \mathbf{y}_T^{T+1}\right)$ gives:

$$\begin{aligned} E\left(y_{T+1}^{T+2} - \alpha_0^* - \boldsymbol{\alpha}^{*\prime}\mathbf{y}_T^{T+1}\right) &= E\left(y_{T+1}^{T+2} - y_{T+1} - \boldsymbol{\alpha}^{*\prime}\left(\mathbf{y}_T^{T+1} - \mathbf{y}_T\right)\right) \\ &= \left(\mu_{v_l} - \sum_{i=1}^{l} \mu_{v_i}\right) + \left(\mu_{\varepsilon_l} - \mu_{\varepsilon_1}\right) - \boldsymbol{\alpha}^{*\prime}\left[\left(\boldsymbol{\mu}_v + \mathbf{i}\mu_{v_l}\right) + \left(\boldsymbol{\mu}_\varepsilon + \mathbf{i}\mu_{\varepsilon_l}\right)\right] \end{aligned}$$

since $E\left(y_{T+1}^{T+2}\right) = \mu_{\widetilde{y}} + E\left(v_{T+1}^{T+2}\right) + E\left(\varepsilon_{T+1}^{T+2}\right) = \mu_{\widetilde{y}} - \sum_{i=1}^{l} \mu_{v_i} - \mu_{\varepsilon_1}$, and $E(y_{T+1}) = \mu_{\widetilde{y}} + E\left(v_{T+1}^{T+1+l}\right) + E\left(\varepsilon_{T+1}^{T+1+l}\right) = \mu_{\widetilde{y}} - \mu_{v_l} - \mu_{\varepsilon_l}$, and similarly for $E\left(\mathbf{y}_T^{T+1}\right)$ and $E(\mathbf{y}_T)$. Recall that $\boldsymbol{\mu}_\varepsilon \equiv E\left([\varepsilon_{t-1}^t, \ldots, \varepsilon_{t-p}^t]\right) = \left[-\mu_{\varepsilon_1}, ..., -\mu_{\varepsilon_p}\right]$, $\boldsymbol{\mu}_v \equiv E\left([v_{t-1}^t, \ldots, v_{t-p}^t]\right) = \left[-\sum_{i=1}^{l} \mu_{v_i}, ..., -\sum_{i=p}^{l} \mu_{v_i}\right]$

and $\mu_{\widetilde{y}} = (1 - \rho(1))^{-1} \left[\rho_0 + \sum_{i=1}^{l} \mu_{v_i}\right]$. The expression for this bias will not equal zero if the mean of the revisions are not zero. ■

B Appendix

B.1 Seasonality of Data Revisions and Statistical Framework

We present an adaptation of the latent news-noise model in section 3 to allow for the seasonal nature of the BEA data revisions. Assuming zero-mean revisions (for simplicity), $\mathbf{v}_t$ and $\boldsymbol{\varepsilon}_t$ defined in (4) need to depend on the quarter of the year that t falls in. Assuming $l = 14$, and that the final estimate y_t^{t+14} reveals the truth, for news we have $\mathbf{v}_{t,t\in Q_j} = -\mathbf{V}_j$, where $j = 1, 2, 3, 4$, and the $\mathbf{V}_j$ matrices are given by:

$$\mathbf{V}_1 = \begin{bmatrix} \sum_{i=1}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ 0 \end{bmatrix}, \quad \mathbf{V}_2 = \begin{bmatrix} \sum_{i=1}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{V}_3 = \begin{bmatrix} \sum_{i=1}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{V}_4 = \begin{bmatrix} \sum_{i=1}^4 v_{i,t} \\ \sum_{i=2}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ \sum_{i=3}^4 v_{i,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ v_{4,t} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

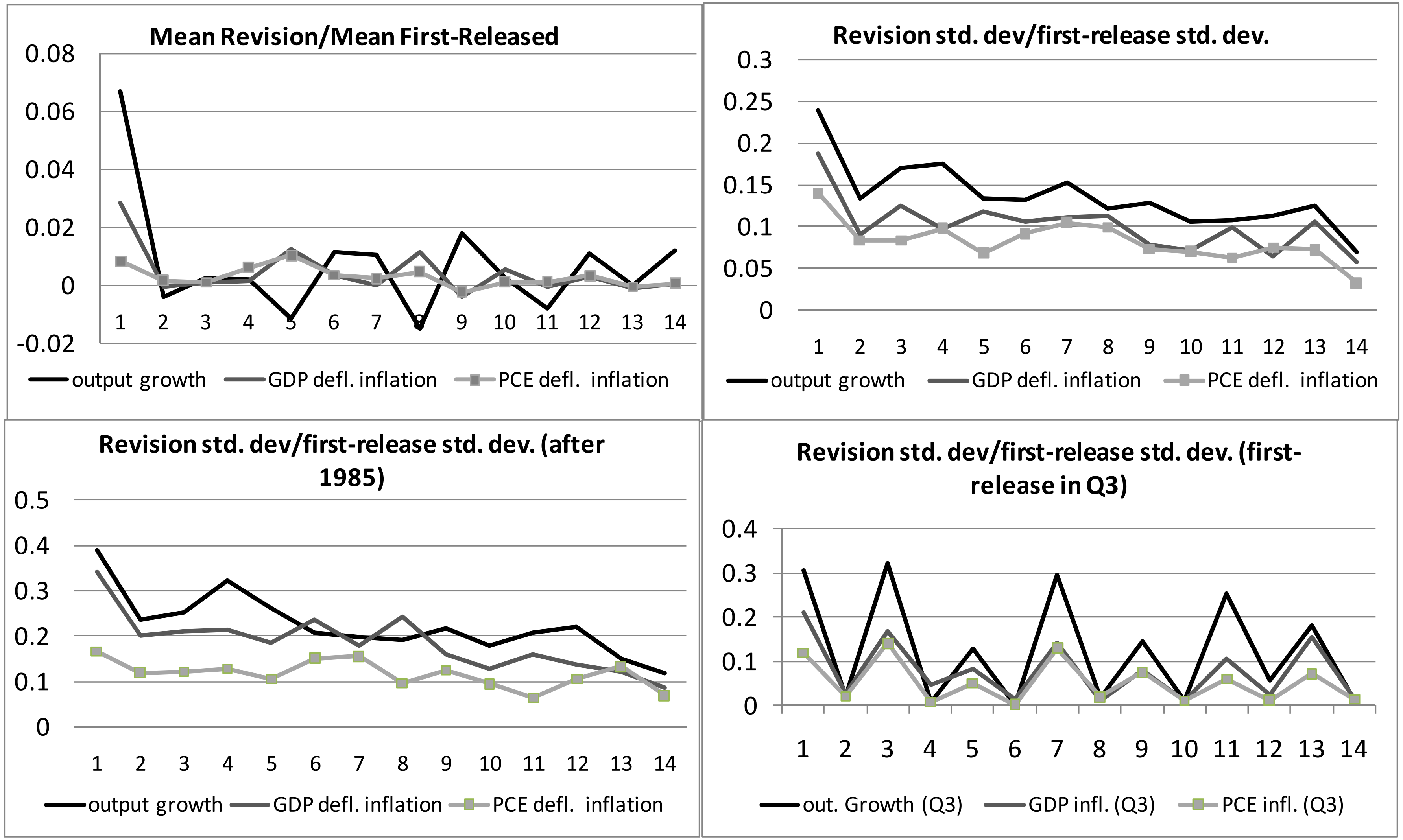
where $v_{i,t} = \sigma_{v_i} \eta_{2t,i}$.

Hence there are only 4 underlying variance terms, so we have assumed, for example, that the variance of the first revision $(y_t^{t+1} - y_t^{t+2})$ does not depend on the quarter. Alternatively one might have $\{v_{1,t}, v_{2,t}, v_{3,t}, v_{4,t}\}$ indexed by quarter.

For the case of noise, the expression for $\boldsymbol{\varepsilon}_t$ is replaced by:

$$\boldsymbol{\varepsilon}_{t,t\in Q_1} = \begin{bmatrix} \sigma_{\varepsilon_1}\eta_{3t,1} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ 0 \end{bmatrix}, \ \boldsymbol{\varepsilon}_{t,t\in Q_2} = \begin{bmatrix} \sigma_{\varepsilon_1}\eta_{3t,1} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ 0 \\ 0 \end{bmatrix}, \ \boldsymbol{\varepsilon}_{t,t\in Q_3} = \begin{bmatrix} \sigma_{\varepsilon_1}\eta_{3t,1} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ \boldsymbol{\varepsilon}_{t,t\in Q_4} = \begin{bmatrix} \sigma_{\varepsilon_1}\eta_{3t,1} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_3}\eta_{3t,3} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ \sigma_{\varepsilon_4}\eta_{3t,4} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Figure 1: The mean and standard deviation of the revisions relative to the first-released data.



Note: Revisions are defined as $r_{i,t} = y_t^{t+1+i} - y_t^{t+i}$ for i = 1,...,l.

Table 1: Characteristics of the data generating processes.

DGP n.	ρ_0	ρ_1	ρ_2	R_1	μ_{r_1}, μ_{r_5}	$\sigma_{y_t^{t+1}}$	$\sigma_{r_1} / \sigma_{y_t^{t+1}}$	$\sigma_{r_2,\ldots,r_{l-1}} / \sigma_{y_t^{t+1}}$	$\sigma_{r_l} / \sigma_{y_t^{t+1}}$
1	.4	.2	.2	.5	.06, .03	.589	.4	.2	.1
2	.4	.2	.2	.5	.06, .03	.589	.6	.3	.15
3	.4	.2	.2	.5	.06, .03	.589	.3	.3	.3
4	.4	.2	.2	.5	.06, .03	.589	.4	.2	0
5	.4	.5	.3	.5	.12, .06	.749	.4	.2	.1
6	.4	.5	.3	.5	.12, .06	.749	.6	.3	.15
7	.4	.5	.3	.5	.12, .06	.749	.3	.3	.3
8	.4	.5	.3	.5	.12, .06	.749	.4	.2	0

Table 2: Analytical values of the estimators, forecast biases and mean squared forecast errors for the DGPs of Table 1 (*l* = *14*)

	AR parameters								Forecasting y_{T+1}^{T+1+1}					
DGP n.	ϕ_0^*	α_0^*	ϕ_1^*	α_1^*	ϕ_2^*	α_2^*	$\phi_1^*+\phi_2^*$	$\alpha_1^*+\alpha_2^*$	Optimal bias	EOS bias	Bias Diff	Optimal MSFE	EOS MSFE	MSFE Ratio
Pure Noise														
1	0.39	0.40	0.17	0.20	0.20	0.20	0.37	0.40	0.000	-0.032	-0.032	0.296	0.298	0.993
2	0.40	0.41	0.15	0.20	0.20	0.20	0.34	0.39	0.000	-0.032	-0.032	0.354	0.356	0.994
3	0.39	0.42	0.19	0.19	0.19	0.19	0.37	0.37	0.000	-0.033	-0.033	0.277	0.278	0.996
4	0.39	0.40	0.17	0.20	0.20	0.20	0.37	0.40	0.000	-0.032	-0.032	0.296	0.298	0.993
5	0.42	0.41	0.39	0.49	0.36	0.30	0.76	0.80	0.000	-0.061	-0.061	0.360	0.367	0.981
6	0.49	0.42	0.32	0.49	0.39	0.30	0.71	0.79	0.000	-0.061	-0.061	0.490	0.508	0.965
7	0.42	0.48	0.46	0.46	0.30	0.30	0.76	0.76	0.000	-0.065	-0.065	0.317	0.321	0.988
8	0.42	0.40	0.39	0.50	0.36	0.30	0.76	0.80	0.000	-0.060	-0.060	0.360	0.368	0.978
Pure News														
1	0.41	0.46	0.22	0.20	0.20	0.20	0.41	0.40	0.000	-0.044	-0.044	0.263	0.263	0.992
2	0.40	0.46	0.24	0.20	0.19	0.20	0.43	0.40	0.000	-0.044	-0.044	0.278	0.280	0.993
3	0.40	0.46	0.24	0.20	0.19	0.20	0.43	0.40	0.000	-0.044	-0.044	0.276	0.279	0.989
4	0.41	0.46	0.22	0.20	0.20	0.20	0.41	0.40	0.000	-0.044	-0.044	0.262	0.264	0.992
5	0.45	0.58	0.60	0.50	0.23	0.30	0.82	0.80	0.000	-0.072	-0.072	0.361	0.371	0.973
6	0.42	0.58	0.68	0.50	0.16	0.30	0.84	0.80	0.000	-0.072	-0.072	0.493	0.516	0.955
7	0.42	0.57	0.69	0.51	0.15	0.29	0.84	0.80	0.000	-0.071	-0.071	0.465	0.489	0.951
8	0.45	0.58	0.60	0.50	0.23	0.30	0.82	0.80	0.000	-0.072	-0.072	0.359	0.369	0.972

Table 3: Small sample biases: AR model parameters

	$\hat{\beta}_0 - \phi_0^*$				$\hat{\alpha}_0 - \alpha_0^*$				$(\hat{\beta}_1 + \hat{\beta}_2) - (\phi_1^* + \phi_2^*)$				$(\hat{\alpha}_1 + \hat{\alpha}_2) - (\alpha_1^* + \alpha_1^*)$			
DGP n.	T=50	T=100	T=200	T=500	T=50	T=100	T=200	T=500	T=50	T=100	T=200	T=500	T=50	T=100	T=200	T=500
Pure Noise																
1	0.06	0.03	0.01	0.01	0.06	0.03	0.01	0.01	-0.09	-0.04	-0.02	-0.01	-0.09	-0.04	-0.02	-0.01
2	0.06	0.03	0.01	0.01	0.06	0.03	0.01	0.01	-0.09	-0.04	-0.02	-0.01	-0.09	-0.05	-0.02	-0.01
3	0.06	0.03	0.01	0.01	0.06	0.03	0.01	0.01	-0.09	-0.04	-0.02	-0.01	-0.09	-0.04	-0.02	-0.01
4	0.06	0.03	0.01	0.01	0.06	0.03	0.01	0.01	-0.09	-0.04	-0.02	-0.01	-0.09	-0.05	-0.02	-0.01
5	0.26	0.12	0.06	0.03	0.24	0.11	0.06	0.02	-0.13	-0.06	-0.03	-0.01	-0.12	-0.06	-0.03	-0.01
6	0.28	0.13	0.07	0.03	0.25	0.12	0.06	0.02	-0.14	-0.07	-0.03	-0.01	-0.13	-0.06	-0.03	-0.01
7	0.25	0.12	0.06	0.02	0.25	0.12	0.06	0.02	-0.13	-0.06	-0.03	-0.01	-0.13	-0.06	-0.03	-0.01
8	0.26	0.12	0.06	0.03	0.24	0.11	0.06	0.02	-0.13	-0.06	-0.03	-0.01	-0.12	-0.06	-0.03	-0.01
Pure News																
1	0.06	0.03	0.01	0.01	0.07	0.03	0.02	0.01	-0.08	-0.04	-0.02	-0.01	-0.09	-0.04	-0.02	-0.01
2	0.05	0.03	0.01	0.01	0.07	0.03	0.02	0.01	-0.07	-0.04	-0.02	-0.01	-0.09	-0.04	-0.02	-0.01
3	0.06	0.03	0.01	0.01	0.07	0.03	0.02	0.01	-0.09	-0.04	-0.02	-0.01	-0.09	-0.05	-0.02	-0.01
4	0.06	0.03	0.01	0.01	0.07	0.03	0.02	0.01	-0.08	-0.04	-0.02	-0.01	-0.09	-0.05	-0.02	-0.01
5	0.27	0.13	0.06	0.03	0.33	0.15	0.08	0.03	-0.10	-0.05	-0.02	-0.01	-0.11	-0.05	-0.03	-0.01
6	0.24	0.11	0.06	0.02	0.32	0.15	0.07	0.03	-0.09	-0.04	-0.02	-0.01	-0.12	-0.05	-0.03	-0.01
7	0.26	0.12	0.06	0.02	0.32	0.15	0.07	0.03	-0.09	-0.04	-0.02	-0.01	-0.11	-0.05	-0.03	-0.01
8	0.27	0.13	0.06	0.03	0.33	0.15	0.08	0.03	-0.10	-0.05	-0.02	-0.01	-0.12	-0.06	-0.03	-0.01

Table 4: Small sample forecast accuracy findings: biases and MSFEs (*l = 14*)

	Forecasting y_{T+1}^{T+1+1}										Forecasting y_{T+4}^{T+4+1}, MSFE ratio			
DGP n.	Bias Diff.	T=50	T=100	T=200	T=500	MSFE Ratio	T=50	T=100	T=200	T=500	T=50	T=100	T=200	T=500
Pure Noise as DGP, Comparing AR_RTV/AR_EOS														
1	*-0.032*	-0.032	-0.021	-0.033	*-0.032*	*0.993*	1.004	0.999	0.997	0.996	1.006	0.998	0.996	0.996
2	*-0.032*	-0.032	-0.020	-0.033	*-0.031*	*0.994*	1.014	0.997	0.996	1.000	1.014	0.999	0.998	0.997
3	*-0.033*	-0.032	-0.023	-0.033	*-0.032*	*0.996*	0.992	1.000	0.995	0.999	0.995	0.995	0.994	0.995
4	*-0.032*	-0.032	-0.021	-0.033	*-0.032*	*0.993*	1.004	0.994	0.997	0.995	1.003	0.997	0.997	0.998
5	*-0.061*	-0.062	-0.051	-0.062	*-0.059*	*0.981*	1.000	0.988	0.983	0.979	1.010	1.004	1.001	1.008
6	*-0.061*	-0.064	-0.050	-0.063	*-0.059*	*0.965*	1.001	0.983	0.977	0.968	1.015	1.009	1.003	1.007
7	*-0.065*	-0.067	-0.057	-0.066	*-0.065*	*0.988*	0.981	0.984	0.986	0.989	0.998	1.001	0.994	0.989
8	*-0.060*	-0.062	-0.050	-0.061	*-0.059*	*0.978*	0.994	0.990	0.983	0.978	0.994	1.009	1.001	1.001
Pure News as DGP, Comparing AR_RTV/AR_EOS														
1	*-0.044*	-0.044	-0.034	-0.043	*-0.044*	0.992	0.976	0.982	0.989	0.994	0.966	0.979	0.983	0.989
2	*-0.044*	-0.044	-0.032	-0.043	*-0.044*	0.992	0.947	0.966	0.982	0.993	0.945	0.969	0.980	0.986
3	*-0.044*	-0.044	-0.033	-0.044	*-0.044*	0.989	0.965	0.978	0.985	0.985	0.955	0.976	0.981	0.989
4	*-0.044*	-0.044	-0.034	-0.043	*-0.044*	0.992	0.976	0.985	0.987	0.996	0.968	0.981	0.986	0.988
5	*-0.072*	-0.073	-0.058	-0.072	*-0.073*	0.973	0.949	0.959	0.962	0.973	0.964	0.976	0.980	0.988
6	*-0.072*	-0.071	-0.055	-0.072	*-0.073*	0.955	0.926	0.941	0.952	0.961	0.958	0.968	0.974	0.982
7	*-0.071*	-0.069	-0.058	-0.071	*-0.071*	0.951	0.925	0.929	0.945	0.947	0.971	0.975	0.981	0.984
8	*-0.072*	-0.073	-0.058	-0.072	*-0.073*	0.972	0.941	0.957	0.966	0.977	0.954	0.981	0.974	0.982
Using VB-VAR as DGP, Comparing AR_RTV/AR_EOS														
9	--	-0.023	-0.033	-0.034	-0.035	--	0.979	0.973	0.977	0.976	0.988	0.988	0.928	0.993
10	--	-0.012	-0.010	-0.011	-0.009	--	0.971	0.959	0.958	0.964	0.976	1.003	1.013	1.023
11	--	0.001	0.004	-0.007	-0.007	--	0.964	0.944	0.947	0.931	0.945	0.983	0.987	1.001
Using VB-VAR as DGP, Comparing VB-VAR/AR_EOS														
9		-0.026	-0.039	-0.032	-0.037		1.427	1.010	0.927	0.892	1.549	0.996	0.921	0.888
10		0.003	-0.009	-0.009	-0.011		1.459	0.950	0.883	0.845	1.767	1.028	0.932	0.881
11		0.001	-0.004	-0.025	-0.006		1.404	0.975	0.893	0.867	1.706	1.089	0.960	0.831

Note: "Bias Diff" are differences between the absolute bias: negative values indicate that RTV (or VB-VAR in last panel) reduces bias. "MSFE ratio" are ratios of MSFE: values smaller than one indicate that RTV (or VB-VAR in last panel) reduces the MSFE. The first columns under "Bias Diff" and "MSFE ratio" report the population analytical values (of Table 2) for ease of comparison. DGPs 9, 10 and 11 are estimated VB-VAR models for, respectively, output growth, GDP deflator inflation and PCE inflation. The autoregressive order of all the AR models is 2, except that AR(4) models are estimated for the VB-VAR inflation DGPs (10 and 11).

Table 5: Characteristics of data releases during the estimation (65-85) and out-of-sample forecast periods (85-06).

		Mean			Standard Deviation			Autocorrelation (1st)			H_0: Mean = 0		H_0: News		H_0:Noise	
		y_t^{t+1}	y_t^{t+14}	y_t^{09Q1}	y_t^{t+1}	y_t^{t+14}	y_t^{09Q1}	y_t^{t+1}	y_t^{t+14}	y_t^{09Q1}	r_{14}	r_{09Q1}	r_{14}	r_{09Q1}	r_{14}	r_{09Q1}
Output growth	1965Q3-1985Q2	0.76	0.79	0.81	1.08	1.07	1.08	0.28	0.29	0.29	[.03]	[.01]	[.08]	[.02]	[.09]	[.01]
	1985Q3-2006:Q4	0.68	0.68	0.75	0.43	0.52	0.50	0.33	0.33	0.23	[.92]	[.06]	[.28]	[.03]	[.00]	[.00]
GDP deflator	1965Q3-1985Q2	1.45	1.43	1.41	0.59	0.59	0.59	0.70	0.75	0.80	[.17]	[.03]	[.48]	[.36]	[.10]	[.03]
	1985Q3-2006:Q4	0.58	0.65	0.60	0.28	0.27	0.23	0.57	0.55	0.56	[.00]	[.30]	[.00]	[.00]	[.00]	[.47]
PCE deflator	1965Q3-1985Q2	1.40	1.39	1.38	0.63	0.64	0.65	0.83	0.83	0.84	[.25]	[.13]	[.52]	[.23]	[.26]	[.04]
	1985Q3-2006:Q4	0.64	0.69	0.64	0.38	0.35	0.30	0.49	0.60	0.56	[.00]	[.76]	[.00]	[.00]	[.00]	[.14]

Note: *p*-values are computed for *F*-statistics (news/noise) and *t*-statistics (mean) using Newey-West standard errors, and are displayed in [].
The revisions are defined as $r_{14} = y_t^{t+14} - y_t^{t+1}$ and $r_{09Q1} = y_t^{09Q1} - y_t^{t+1}$.

Table 6: Comparing the RMSFEs from RTV and EOS in a recursive out-of-sample forecasting exercise (for the period 1985:Q3-2008:Q4; *n = 94* quarters).

6A. Forecasting output growth

	h = 1			h = 4		
	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(1)	**0.976**	0.981	0.999	**0.956**	**0.956**	**0.972**
		0.998	0.995		0.982	0.980
VB-VAR(1), q=4	0.999	0.996	1.013	**0.965**	**0.965**	0.988
VB-VAR(1), q=5	0.995	0.996	1.012	**0.967**	**0.960**	0.980
VB-VAR(1), q=8	0.994	0.986	1.009	1.002	**0.958**	0.989
VB-VAR(1), q=14	1.028	1.026	1.034	1.017	**0.970**	0.990

6B. Forecasting GDP deflator inflation

	h = 1			h = 4		
	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	**0.961**	**0.972**	**0.959**	**0.959**	**0.967**	**0.953**
		0.973	0.977		0.994	0.991
VB-VAR(1), q=4	**0.962**	**0.962**	**0.958**	**0.968**	**0.966**	**0.958**
VB-VAR(1), q=5	0.988	**0.966**	**0.965**	1.034	1.028	1.021
VB-VAR(1), q=8	**0.963**	**0.956**	**0.955**	1.036	1.040	1.027
VB-VAR(1), q=14	**0.955**	**0.953**	**0.959**	**0.973**	**0.978**	**0.970**

6C. Forecasting PCE deflator inflation

	h = 1			h = 4		
	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}
AR(4)	**0.979**	0.982	**0.977**	**0.977**	**0.973**	**0.969**
		0.987	0.988		0.985	0.986
VB-VAR(1), q=4	**0.974**	**0.978**	**0.971**	**0.968**	**0.963**	**0.959**
VB-VAR(1), q=5	1.000	1.005	0.999	1.017	1.018	1.010
VB-VAR(1), q=8	1.016	1.024	1.014	1.049	1.055	1.044
VB-VAR(1), q=14	1.068	1.084	1.080	1.054	1.067	1.052

Note: Entries with values smaller than one indicate that RTV reduces the RMSFE relative to EOS. For VB-VAR models, ratios are relative to the AR(p) with EOS data (same benchmark model for each panel). Entries in the second line are for RTV bias-corrected forecasts (relative to EOS). The correction is based on the difference between the unconditional means of y_t^{t+14} and y_t^{t+1} with data up to the forecast origin. Emboldened entries indicate RMSFE reductions larger than 2%. Multiple step-ahead forecasts are computed by iteration. Models are estimated with increasing windows of data. In-sample period starts in 1959:Q1 for AR comparisons and 1965:Q3 when comparing VAR models with AR with EOS data.

Table 7: Comparison of EOS and RTV data for forecasting output growth and inflation with ADL models of activity variables (for the period 1985:Q3-2008:Q4; *n = 94* quarters).

7A. Forecasting output growth

		h = 0				h = 1				h = 4			
					Ratio AR(1)_RTV				Ratio AR(1)_RTV				Ratio AR(1)_RTV
Model:	sample starts:	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
		With industrial production											
ADL (1,2)	1959:Q1	1.025	1.004	1.009	0.855	1.052	1.038	1.052	1.027	0.972	0.975	0.997	1.052
ADL (1,2)	1979:Q1	**0.921**	**0.958**	**0.928**	**0.850**	**0.971**	0.980	0.981	0.990	0.936	0.964	0.982	1.001
		With employment											
ADL (1,2)	1959:Q1	0.983	1.013	1.011	0.893	1.002	1.043	1.045	1.042	1.014	1.013	1.024	1.131
ADL (1,2)	1979:Q1	**0.921**	**0.958**	**0.968**	**0.890**	0.898	0.975	0.959	1.024	0.979	0.987	1.007	1.023

7B. Forecasting GDP deflator inflation

		h = 0				h = 1				h = 4			
					Ratio AR(4)_RTV				Ratio AR(4)_RTV				Ratio AR(4)_RTV
Model:	sample starts:	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
		With industrial production											
ADL (4,2)	1959:Q1	0.948	0.945	0.945	1.001	**0.945**	**0.958**	**0.943**	**0.985**	1.052	1.006	0.983	1.044
ADL (4,2)	1979:Q1	0.935	0.911	0.940	0.992	0.915	0.899	0.918	1.007	0.946	0.954	0.922	1.032
		With employment											
ADL (4,2)	1959:Q1	0.952	0.974	0.967	1.027	0.968	0.986	0.968	1.016	0.990	1.017	0.983	1.103
ADL (4,2)	1979:Q1	0.940	0.905	0.945	0.996	**0.939**	**0.913**	**0.926**	**0.999**	0.965	0.976	0.948	1.063
		With output growth											
ADL (4,2)	1959:Q1	0.949	0.944	0.939	0.997	**0.953**	**0.974**	**0.953**	**0.979**	**0.962**	1.037	**0.972**	**0.968**
ADL (4,2)	1979:Q1	0.955	0.926	0.958	1.013	0.933	0.911	0.928	1.015	0.958	0.979	0.943	1.049

7C. Forecasting PCE deflator inflation

		h = 0				h = 1				h = 4			
					Ratio AR(4)_RTV				Ratio AR(4)_RTV				Ratio AR(4)_RTV
Model:	sample starts:	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+1}^{T+1+1}	y_{T+1}^{T+15}	y_{T+1}^{09Q1}	y_{T+1}^{09Q1}	y_{T+4}^{T+4+1}	y_{T+4}^{T+4+14}	y_{T+4}^{09Q1}	y_{T+4}^{09Q1}
		With industrial production											
ADL (4,2)	1959:Q1	**0.977**	**0.979**	**0.976**	**0.978**	**0.973**	**0.979**	**0.974**	**0.984**	1.020	1.019	1.013	1.030
ADL (4,2)	1979:Q1	0.988	1.000	0.981	0.987	**0.979**	0.987	**0.977**	**0.994**	0.983	0.967	0.968	1.016
		With employment											
ADL (4,2)	1959:Q1	**0.974**	**0.979**	**0.973**	**0.968**	**0.973**	**0.981**	**0.974**	**0.983**	1.015	1.020	1.013	1.006
ADL (4,2)	1979:Q1	0.988	0.984	0.982	0.993	**0.975**	0.980	**0.974**	**0.995**	0.973	0.977	0.964	1.007
		With output growth											
ADL (4,2)	1959:Q1	0.986	0.991	0.993	1.013	**0.974**	0.986	**0.972**	**0.991**	1.039	1.052	1.031	1.010
ADL (4,2)	1979:Q1	0.997	1.000	0.996	1.007	0.933	0.911	0.928	1.015	0.991	0.994	0.973	1.012

Note: Entries with values smaller than one indicate that RTV reduces the RMSFE relative to EOS, except entries in the column "Ratio AR(p)_RTV" where values smaller than one indicate that the ADL model reduces the RMSFE relative to the AR model. ADL models are estimated for each forecast horizon (specification as described in the first column) and forecasts are computed `directly' with increasing windows of data. As in the previous tables, the variable we aim to forecast is the quarterly difference at an annualized rate *h*-quarters ahead, namely, $y_{t+h} = 400[\ln(X_{t+h}) - \ln(X_{t-h-1})]$, where X is the level of real GDP, the GDP deflator or the PCE deflator. The explanatory variables are also annualized quarterly differences. We use as our quarterly vintage of industrial production and employment the vintage published in the first month of each quarter, and calculate a quarterly series by averaging the monthly observations, before computing first differences. When estimating with RTV data, the vintage of the RHS variables depends on the forecasting horizon as described in the text.