Improving real-time estimates of output gaps and inflation trends with multiple-vintage VAR models.

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

February 9, 2011

Abstract

Real-time estimates of output gaps and inflation trends differ from the values that are obtained using data available long after the event. Part of the problem with real-time estimates is that the data on which the estimates are based is subsequently revised. We show that vector-autoregressive models of data vintages can provide forecasts of post-revision values of future observations and of already-released past observations capable of improving estimates of output gap and inflation trend in real time. The fact that revisions to output and inflation data are predictable, and that past data vintages help to forecast future estimates, accounts for the improvements achievable in real time.

Keywords: data revisions, real-time forecasting, output gap, inflation trend.

JEL code: C53.

Improving real-time estimates of output gaps and inflation trends with multiple-vintage VAR models.

February 9, 2011

Abstract

Real-time estimates of output gaps and inflation trends differ from the values that are obtained using data available long after the event. Part of the problem with real-time estimates is that the data on which the estimates are based is subsequently revised. We show that vector-autoregressive models of data vintages can provide forecasts of post-revision values of future observations and of already-released past observations capable of improving estimates of output gap and inflation trend in real time. The fact that revisions to output and inflation data are predictable, and that past data vintages help to forecast future estimates, accounts for the improvements achievable in real time.

Keywords: data revisions, real-time forecasting, output gap, inflation trend.

JEL code: C53.

1   Introduction

Policy makers often wish to have reliable estimates of the true current state of the economy. Because national accounts data are subject to revision, post-revision estimates of the value of key macroeconomic variables (such as output growth and inflation) in the current and recent quarters will not be available until a number of years in the future, even if we are prepared to assume that the data are subject to a finite revision process. Yet estimates of these figures are typically required today to guide government policy. This suggests the use of models to forecast the 'true' or revised values of these macro-variables.

Monetary policy perhaps best exemplifies the need to estimate the current state of the economy. Estimates of the current output gap are key, and it has been shown that estimates based on final data can be markedly different from those available in real time, affecting both historical evaluations of monetary policy and the effective conduct of monetary policy in real time (see, e.g., Orphanides (2001) and Orphanides and van Norden (2002)). We consider the estimation of output gaps in real-time. Models are employed to calculate post-revision estimates of the level of output for current, past and future observations. These estimates are then used to improve real-time estimates of the output gap. The output gap estimates so obtained exhibit higher correlation with the 'true' gap than standard real-time estimates.

Recently, a number of authors[1] have estimated the trend inflation rate, associated with either long-run inflation expectations or a time-varying inflation target. Changes in the persistence of the inflation gap, defined as the difference between the observed inflation rate and the trend inflation, are important for the conduct of monetary policy (see Cogley *et al.* (2010)). The literature to date has used fully-revised data: our interest is in the quality of estimates that can be obtained in real time, and as for the output gap, whether we can obtain more accurate estimates by predicting the future revisions to the relevant data. We find that the real-time estimates of trend inflation (measured by either the GDP or PCE deflators) are significantly lower than estimates with fully-revised data for the 1995-2005 period. However, we show that the bias in these real-time estimates can be removed if we use model predictions of data revisions.

The models we use to forecast revisions and post-revision future observations are related to the vector-autoregressive (VAR) models of Garratt, Lee, Mise and Shields (2008, 2009) and Hecq and Jacobs (2009). In this context, the elements of the vector of variables being modelled consist of the

---

[1]These include contributions by Kozicki and Tinsley (2005), Stock and Watson (2007, 2010) and Cogley, Primiceri and Sargent (2010).

vintage-$t+1$ estimates of the observations for the period t back to $t-q+1$ (for either output growth, or inflation). The model supposes that these vintage-$t+1$ estimates are predictable from earlier vintage estimates (e.g., the vintage-t estimate of observations $t-1$ to $t-q$). We consider a number of related VAR models, including a model that more closely reflect the timing of revisions to data by the government statistical agency. Such a model recognizes that the US national accounts revisions process is seasonal, in the sense that the pattern of revisions to a particular observation depends on the quarter of the year to which the observation belongs. The literature on the VAR modelling of data revisions largely neglects this characteristic.[2] The US Bureau of Economic Analysis (BEA) publishes annual revisions to national accounts data every July, which affect the three years of data published up to that point. One of our points will be whether the standard practice of disregarding the seasonal nature of these regular annual revisions is a harmful simplification, in terms of forecasting future observations and data revisions. On the face of it, one might expect that specifying the VAR model to more closely mirror the operating procedures of the BEA would improve the model's performance in improving real-time estimates of the output gap and inflation trend. Another model imposes restrictions based on the widely-held belief that estimates after the first revision are all but unpredictable.[3] We assess the impact of this assumption about the nature of BEA revisions on forecasting performance and subsequently on the real-time measurement of the output gap and inflation trend.

The plan of the remainder of the paper is as follows. Section 2 describes the basic VAR model and the alternative versions. Section 3 provides evidence on the extent to which data revisions are predictable using vintage-based VAR models, after first describing the in-sample fit of the different models. This section can also be viewed as a precursor to sections 4 and 5, where the ability of the models to forecast revisions and future estimates of output growth and inflation is shown to translate into improved measures of the output gap and the inflation trend and gap, respectively. Section 6 offers some concluding remarks.

---

[2] Kishor and Koenig (2010) suggest an extension to their modelling approach that considers the periodicity of data revisions in an appendix. However they do not apply the model to data.

[3] See, for example, Garratt *et al.* (2008) and Clark (2010). Both use the BEA 'final' estimate, available two quarters later, as actual values for computing forecast errors. The BEA 'final' estimate corresponds to our second estimate (equivalently, once revised value). Their choice of target variable is based on the assumption that annual and benchmark revisions are largely unpredictable.

2 The Vintage-based VARs

Our basic (non-seasonal) model is closely related to the vintage-based VAR (V-VAR) of Hecq and Jacobs (2009) and the models of Garratt *et al.* (2008, 2009). We briefly describe this model and how it relates to other models in the literature, before explaining how it can be modified to better capture the nature of US data revisions. The models we consider retain the benefits of modelling relationships between observable variables without the need to introduce unobserved components. Examples of models of data revisions with unobserved components include Jacobs and van Norden (2007), which relates a vector of different vintage estimates of y_t to the (generally unobserved) true value and latent news and noise measurement errors, and Cunningham, Eklund, Jeffery, Kapetanios and Labhard (2009), *inter alia*. Kishor and Koenig (2010) (building on earlier contributions by Howrey (1978, 1984) and Sargent (1989)) estimate a VAR 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.

We work with growth rates,[4] defined either by quarterly changes $y_t^{t+1} = 400 \left(\ln Y_t^{t+1} - \ln Y_{t-1}^{t+1} \right)$ or by annual changes $y_t^{t+1} = 100 \left(\ln Y_t^{t+1} - \ln Y_{t-4}^{t+1} \right)$, computed using data from vintage $t+1$. The superscript on y denotes the data vintage, and the subscript the time period to which the observation refers. Hence y_t^{t+1} is the first estimate of the growth rate observed at t. If we suppose that there are revisions for the next $q-1$ quarters, but thereafter the observation is unrevised (i.e., $y_t^{t+q+i} = y_t^{t+q}$ for $i > 0$) then we can model the vintage $t+1$ values of observations $t-q+1$ through t as a V-VAR:

$$\mathbf{y}^{t+1} = \mathbf{c} + \sum_{i=1}^{p} \mathbf{\Gamma}_i \mathbf{y}^{t+1-i} + \boldsymbol{\varepsilon}^{t+1} \tag{1}$$

where $\mathbf{y}^{t+1} = \left[y_t^{t+1}, y_{t-1}^{t+1}, \ldots, y_{t-q+1}^{t+1} \right]'$, $\mathbf{y}^{t+1-i} = \left[y_{t-i}^{t+1-i}, y_{t-1-i}^{t+1-i}, \ldots, y_{t-q+1-i}^{t+1-i} \right]'$, and $\mathbf{c}$ is $q \times 1$, $\boldsymbol{\varepsilon}^{t+1}$ is $q \times 1$. Notice that y_{t-q}^{t+1} would be redundant if added as the $(q+1)^{th}$ element of $\mathbf{y}^{t+1}$, because y_{t-q}^{t+1} is identical to y_{t-q}^t (the last element of $\mathbf{y}^t$). The V-VAR models the dynamics of

---

[4]Following Patterson (1995), Garratt *et al.* (2008, 2009) work in terms of the level (of the log) of output, so that consideration of issues to do with whether different vintages are cointegrated arise. They assume the revisions are stationary. Instead, we work with growth rates, and revisions in growth rates. On the face of it, ignoring long-run information constitutes a form of model mis-specification. That said, there is now a body of literature that suggests that error-correcting models will perform poorly when there are (unmodelled) shifts (see, e.g., Clements and Hendry (2006)) as in the case of shifts caused by occasional changes in the base year or other methodological changes in the definition of the series. We regard models specified solely in terms of growth rates as likely to be more robust to the effects of changes in levels due to re-basing and other benchmark revisions. Clements and Galvão (2011) compare the V-VAR model's forecasting performance against that of Garratt *et al.* (2008).

successive vintages of data that include both a new observation y_t^{t+1} and revised estimates of past observations $y_{t-1}^{t+1}, \ldots, y_{t-q+1}^{t+1}$. The the first equation models first releases, y_t^{t+1}, the second equation data that has been revised once, y_{t-1}^{t+1}, and so on. The variance-covariance matrix of the disturbances ($\Sigma_\varepsilon = E(\boldsymbol{\varepsilon}^{t+1}\boldsymbol{\varepsilon}^{t+1\prime})$) captures the correlations between data revisions published in the same vintage.

The autoregressive order p is set to capture serial correlation in $\mathbf{y}^{t+1}$, and in practice it may be possible to set p to a low value, as we now explain. Suppose that there are no data revisions, so $y_{t-1}^{t+1} = y_{t-1}^t$, $y_{t-2}^{t+1} = y_{t-2}^t$, etc., and the superscript is superfluous. As a consequence, the model would collapse to a single AR(q) for y_t ($\equiv y_t^{t+1}$), say $\gamma(L) y_t = \varepsilon_t$, where $\gamma(L) = 1 - \gamma_1 L - \gamma_2 L^2 - \ldots - \gamma_q L^q$, where in terms of (1), $\gamma_1 = \Gamma_{1,11}$, $\gamma_2 = \Gamma_{1,12}$, to $\gamma_q = \Gamma_{1,1q}$, where $\Gamma_{i,jk}$ is the row j, column k element of Γ_i. More generally, if we allow for data revisions, it is clear that the first-order VAR allows for longer lags and more complex dynamic interaction than the first-order nature of the model might suggest. For this reason, low values of p might often suffice to capture the dynamics of $\mathbf{y}^{t+1}$.

The V-VAR model necessarily limits the number of revisions that are allowed - whereas ideally the vector $\mathbf{y}^{t+1}$ should include all the revisions. In practice y_t^{t+q} will not be the same as the value for y_t in the latest vintage of data available to the investigator (perhaps primarily because of benchmark revisions), although it is less clear what the implications of the remaining data uncertainty will be. In the literature, Jacobs and van Norden (2007) allow that the true value may not equal the last available vintage in their unobserved components model, but in the VAR setup this issue is not so easily handled. There is an incentive to consider a large number of data revisions in order that y_t^{t+q} is as near as possible to the value in the last available vintage, leading to high-dimensional VARs with a high degree of parameter estimation uncertainty. We consider a restricted V-VAR model (section 2.2) that helps to counter the proliferation of parameters as q is increased.

The V-VAR is also closely related to some of the single-equation approaches in the literature (such as e.g., Koenig, Dolmas and Piger (2003) and Clements and Galvão (2010)). The first equation of (1) with $p = 1$ is the real-time-vintage autoregression of Clements and Galvão (2010):

$$y_t^{t+1} = \beta_0 + \sum_{i=1}^{q} \beta_i y_{t-i}^t + \varepsilon_t.$$

When estimated on $t = 2, \ldots, T$, and used to forecast y_{T+1}^{T+2}, this approach is shown to be preferable in some circumstances to the traditional or 'end-of-sample' approach. Because the VAR is estimated

equation-by-equation by OLS, the estimates of the coefficients in the first equation (first row) of (1) will be the same as the AR estimated with real-time-vintage data when the orders of the models match. The advantage of the systems-based VAR over the single-equation approach of Clements and Galvão (2010) is that we are able to forecast vintages other than the first estimate (of the observation of interest). In the empirical exercise in section 3, the benchmark to predict future observations is a standard autoregression estimated solely using the latest vintage of data available at the time the forecast is made (that is, using the traditional real-time approach).

2.1 Periodic specifications

The V-VAR ignores the periodicity of the publication of data revisions. In the case of US BEA data, the elements of $\mathbf{y}^{t+1}$ other than the first two (i.e., y_t^{t+1}, the 'advance' estimate, and y_{t-1}^{t+1}, the 'final' estimate of y_{t-1}) will typically remain unchanged unless the $t+1$-vintage is an annual $(t+1 \in Q3)$ or a benchmark revision.[5] In practice, a small modification to this simple seasonal pattern is required. When benchmark revisions are anticipated to be published in January, annual revisions may not be published in the previous July. As a consequence, we define two dummy variables: $D_1^{t+1} = 1$ if an annual revision has been published (always in the third quarter), and $D_2^{t+1} = 1$ if either an annual or a benchmark revision has been published. The majority of cases that $D_2^{t+1} = 1$ are for vintages published in the third quarter, but there are instances for other quarters, especially the first quarter. When $s = 1$, so that only annual revisions are included, the model is known as a seasonal vintage-based VAR (SV-VAR). When $s = 2$, so that both annual and benchmark revisions are included, the model is termed a seasonal and benchmark vintage-based VAR (SBV-VAR).

The model when $p = 1$ is:

$$\mathbf{y}^{t+1} = \left[\widetilde{\mathbf{c}} + \widetilde{\mathbf{\Gamma}}_1 \mathbf{y}^t \right] (1 - D_s^{t+1}) + \left[\mathbf{c} + \mathbf{\Gamma}_1 \mathbf{y}^t \right] D_s^{t+1} + \mathbf{v}^{t+1} \tag{2}$$

where:

$$\widetilde{\Gamma}_1 = \begin{bmatrix} & \boldsymbol{\gamma}_{2 \times q} & \\ \mathbf{0}_{(q-2) \times 1} & \mathbf{I}_{(q-2) \times (q-2)} & \mathbf{0}_{(q-2) \times 1} \end{bmatrix} \tag{3}$$

and $\widetilde{\mathbf{c}} = (c_1, c_2, 0, ..., 0)'$.

Thus, when quarter $t+1$ does not incorporate an annual or benchmark revision, then $D_s^{t+1} = 0$,

[5]The GNP/GDP data of the BEA are subject to three annual revisions in the July of each year: see, e.g., Fixler and Grimm (2005, 2008) and Landefeld, Seskin and Fraumeni (2008).

and hence $y_{t+1-i}^{t+1} = y_{t+1-i}^{t}$ (up to a random error term, v_{t+1-i}^{t+1}) for $i = 3, \ldots, q$. But when $D_s^{t+1} = 1$, y_{t+1-i}^{t+1} is determined by the coefficients in the ith row of $\mathbf{\Gamma}_1$ multiplied into $\mathbf{y}^t$. Hence the above model captures the seasonal aspect of the BEA revisions to national accounts data because whether, say, $y_t^{t+3} - y_t^{t+2} = 0$ depends on the quarter of the year to which t belongs ($y_t^{t+3} - y_t^{t+2} = 0$ unless t falls in Q4).

By way of contrast, the V-VAR simply estimates (2) with $D_s^{t+1} = 1$ for all t: this means that the estimate of $\mathbf{\Gamma}_1$ in (1) will be an average of observations for which $y_{t+1-i}^{t+1} = y_{t+1-i}^{t}$, $i \geq 3$ (annual/benchmark revisions are not published) and observations for which $y_{t+1-i}^{t+1} \neq y_{t+1-i}^{t}$, $i \geq 3$. In theory, we might also let the publication of an annual revision affect the first and second equations, viz, y_t^{t+1} and y_{t-1}^{t+1}, but we assume that the first two rows of $\mathbf{\Gamma}_1$ and $\widetilde{\mathbf{\Gamma}}_1$ are equal to keep the model relatively simple. Finally, note that the form of $\widetilde{\mathbf{c}}$ is such that intercepts are only estimated in the equations for $(y_{t-2}^{t+1}, \ldots, y_{t-q+1}^{t+1})$ when $D_s^{t+1} = 1$.

In Clements and Galvão (2011), we consider additional specifications for modelling annual and benchmark revisions. These do not significantly improve upon the SV-VAR and the SBV-VAR in terms of forecast performance, and hence are not considered here.

2.2 A restricted specification

The V-VAR can be restricted by imposing the condition that, after a (relatively small) number of revisions, further revisions are unpredictable. Suppose for instance that after $n - 1$ revisions, the next estimate y_t^{t+n+1} is an efficient forecast in the sense that the revision from y_t^{t+n} to y_t^{t+n+1} is uncorrelated with y_t^{t+n}, i.e., $E\left[\left(y_t^{t+n+1} - y_t^{t+n}\right) \mid \mathbf{y}^{t+n}\right] = 0$, whereas $E\left[\left(y_t^{t+i+1} - y_t^{t+i}\right) \mid \mathbf{y}^{t+i}\right] \neq 0$ for $i < n$. We can impose this restriction on the VAR, where it translates to $E\left(y_{t-n}^{t+1} \mid y_{t-n}^{t}\right) = y_{t-n}^{t}$. This is achieved by specifying Γ_1 in (1) as:

$$\widetilde{\Gamma}_1 = \begin{bmatrix} & \boldsymbol{\gamma}_{n \times q} & \\ \mathbf{0}_{(q-n) \times (n-1)} & \mathbf{I}_{(q-n) \times (q-n)} & \mathbf{0}_{(q-n) \times 1} \end{bmatrix} \tag{4}$$

and when $p > 1$, it requires in addition that the Γ_i coefficient matrices are restricted to:

$$\widetilde{\Gamma}_i = \begin{bmatrix} \boldsymbol{\gamma}_{i, n \times q} \\ \mathbf{0} \end{bmatrix}$$

for all $i > 1$.

In the empirical work, we set $n = 2$, so that estimates after the first revision are assumed to

be efficient forecasts (i.e., the BEA estimate published two quarters after the period to which it refers is an efficient forecast). Compared to the periodic models, the restricted VAR imposes the restriction that $\Gamma_1 = \widetilde{\Gamma}_1$ (and $\Gamma_i = \widetilde{\Gamma}_i$ when $p > 1$) for all t. The restricted VAR (for $p = 1$ and $n = 2$) is:

$$\mathbf{y}^{t+1} = \mathbf{c} + \widetilde{\mathbf{\Gamma}}_1 \mathbf{y}^t + \mathbf{v}^{t+1} \tag{5}$$

where $\widetilde{\mathbf{\Gamma}}_1$ is given by (3). An unrestricted intercept is included in each equation, so the interpretation of the model is that revisions $y_{t+1-i}^{t+1} - y_{t+1-i}^t$ are uncorrelated with y_{t+1-i}^t for $i \geq 3$, but may be non-zero mean. For example, $E\left(y_{t+1-i}^{t+1} - y_{t+1-i}^t\right) = c_i$, for $i \geq 3$, where c_i is the ith element of $\mathbf{c}$. We name this model the 'news-restricted' vintage-based VAR, RV-VAR.

We describe our approach to the estimation of the VAR models in the appendix.

3   Forecasting with vintage-based VAR models

Section 3.1 describes the in-sample fit of the models, and in particular, the effect of modelling the periodic nature of the publication of data revisions. In Section 3.2 we assess the extent to which the improved fit translates into a superior out-of-sample forecast performance. Our interest is in forecasting the final or post-revision values of both recent and past observations and of future observations, as it is the post-revision values that are required for calculating real-time estimates of the output gap and the inflation trend-gap decomposition in sections 4 and 5.

3.1   In-sample Fit

The models we consider are the unrestricted vintage-based VAR (V-VAR), the specifications that consider the impact of annual and benchmark revisions (SV-VAR and SBV-VAR) and the model that assumes that data revisions published after the 'final' estimate, y_t^{t+2}, are unpredictable except that they may have a non-zero mean (RV-VAR). The RV-VAR model is nested within the other models, so that formal likelihood ratio tests of the other models against the RV-VAR are applicable, but note that all the models other than the RV-VAR have the same number of parameters.[6]

The real-time data on real output and the GDP deflator are published by the Philadelphia

---

[6]Assuming that $q = 14$, the RV-VAR contains 42 parameters: 14 slope parameters in the equations for y_t^{t+1} and y_{t-1}^{t+1}, and an intercept in each of the 14 equations. The V-VAR contains 14×14 slope parameters, and 14 intercepts. The S(B)V-VAR is the same but the slope and intercepts are set to zero in the equations for $y_{t-2}^{t+1}, \ldots, y_{t-q+1}^{t+1}$ when $D_s^{t+1} = 0$.

Fed (see Croushore and Stark (2001)). There are 8 benchmark revisions in the data vintages from 1965:Q4 up 2010:Q1 that comprise our estimation sample. Because the BEA skips the annual (Q3) revision when a benchmark revision is to be published in the first quarter of the following year, there are 36 annual Q3 revisions ($D_1^{t+1} = 1$) rather than the 44 that would otherwise have occurred. There are 44 combined benchmark and annual revisions ($D_2^{t+1} = 1$) - the 8 benchmark revisions and the 36 annual revisions.

Table 1 summarizes information on the fits of these models. We set $p = 1$ and $q = 14$, and the data are expressed in quarterly differences for both real output and the GDP deflator. The choice of $q = 14$ ensures that the data will have undergone the three annual revisions irrespective of the quarter in which it was first published. All models are estimated with vintages from 1965:Q4 up 2010:Q1 (178 vintages). From table 1, the SBV-VAR has the best fit, while the LR statistics reject the restrictions imposed by the 'news-restricted' RV-VAR specification against all the alternative models. This suggests some predictability of annual and benchmark revisions, because the difference between the RV-VAR and the other models is that the RV-VAR assumes estimates after the first revision (y_{t-1-i}^{t+1}, $i > 0$) are unpredictable from $\mathbf{y}^t$, $\mathbf{y}^{t-1}$,.... However, comparisons of likelihoods across models are not necessarily informative about the out-of-sample forecasting performance of the models.

3.2 Empirical Forecasting Comparisons

Consider a forecast origin $T+1$. At this time, all the data vintages up to and including the time-$T+1$ vintage are known, i.e., $\mathbf{y}^{T+1}$, $\mathbf{y}^T$, ... where $\mathbf{y}^{T+1-i} = (y_{T-i}^{T+1-i}, ..., y_{T-i-q+1}^{T+1-i})'$ for $i = 0, 1, 2, \ldots$. The h-step ahead forecast of the vector $\mathbf{y}^{T+1+h}$ is defined as $\mathbf{y}^{T+1+h|T+1} \equiv E\left(\mathbf{y}^{T+1+h} \mid \mathbf{y}^{T+1}, \mathbf{y}^T, \ldots\right)$. The elements of $\mathbf{y}^{T+1+h|T+1}$ are $(y_{T+h}^{T+1+h|T+1}, ..., y_{T+h-q+1}^{T+1+h|T+1})$, and thus provide forecasts of the first estimate of y_{T+h}, a forecast of the second estimate of y_{T+h-1}, and so on down to the qth estimate of $y_{T+h-q+1}$. The forecasts are computed by iteration in the standard way based on the estimated models.

Clements and Galvão (2011) consider all these elements of the forecast vector $\mathbf{y}^{T+1+h|T+1}$, and evaluate these forecasts against: i) the maturities of the data that are explicitly being forecast (so, e.g., the jth element $\mathbf{y}_{T+1+h-j}^{T+1+h|T+1}$ is compared to the actual value $\mathbf{y}_{T+1+h-j}^{T+1+h}$), and ii) the last-available estimates of the actual values (i.e., $\mathbf{y}_{T+1+h-j}^{T+1+h|T+1}$ is compared to the actual value $\mathbf{y}_{T+1+h-j}^{2010:1}$). Our two empirical applications require forecasts of 'post-revision' values, which are taken to be the $q = 14$ estimates. This means we only consider forecasts of a subset of the elements of $\mathbf{y}^{T+1+h|T+1}$,

namely, the last elements of each vector for $h = 1, 2, 3, \ldots, h^*$. For forecast origin $T + 1$, this gives the following set of forecasts: $\mathbf{y}_{T+2-q}^{T+2|T+1}$, $\mathbf{y}_{T+3-q}^{T+3|T+1}$, $\ldots$, $\mathbf{y}_{T+1+h^*-q}^{T+1+h^*|T+1}$. Hence when $\tau \equiv 1 + h - q = 0$, we have a forecast of the post-revision value of the present (y_T), and for $\tau < 0$ forecasts of post-revision past data $(y_{T-1}, y_{T-2}, \text{etc})$, and for $\tau > 0$ forecasts of future revised observations $(y_{T+1}, y_{T+2}, \text{etc.})$. Thus, short-horizon forecasts (e.g., $h = 1$) will provide forecasts of (distant) past data $(\tau < 0)$, and we will obtain forecasts of more recent past data, and then future observations, as h increases.

We compute individual MSFE-measures of accuracy of forecasts $\hat{\mathbf{y}}_{T+1+h-q}^{T+1+h|T+1}$ with $q = 14$, by averaging over the out-of-sample forecast origins from $T+1$ up to $T+N$, for each of $h = 1, 2, \ldots, h^*$. As summary measures, we also sum the the MSFEs over h for $\tau \leq 0$, corresponding to revisions to past data, and for $\tau > 0$, to emphasize any differences in forecast performance between forecasting 'future observations' and forecasting 'data revisions'.

The assessment of the value of the VAR model forecasts is made relative to some simple benchmark forecasts. For future observations, we compute forecasts using an $\text{AR}(p)$ model estimated on the data vintage available at the forecast origin, that is, the end-of-sample approach to estimate autoregressive models in real time. For data revisions - future data releases of past observations - the benchmark is such that revisions are assumed to be zero. Hence the 'no-change predictor' at forecast origin $T + 1$ is $\mathbf{y}_{T+1+h-q}^{T+1+h|T+1} = \mathbf{y}_{T+1+h-q}^{T+1}$ for $\tau \leq 0$.

By comparing the accuracy of the forecasts of data revisions from the VAR models with the forecasts from the no-change benchmark we can assess the predictability of data revisions. In particular, the RV-VAR imposes the restriction that revisions to data that have already been revised once are unpredictable, apart from being systematically either upward or downward, so would be expected to do better than the benchmark if revisions are non-zero mean. If the V-VAR outperforms the RV-VAR, then the annual and benchmark revisions would appear to be in part predictable. In terms of forecasting future observations, we are interested in whether the use of multiple data vintages, as in the V-VAR models, leads to superior forecasts relative to forecasting with just the latest data vintage using an AR model.

We initiallly estimate the VAR models on the data vintages from 1965:Q4 to 1995:Q3, and then on vintages from 1965:Q4 to 1995:Q4, and so on up to 1965:Q4 to 2006:Q3, adopting an expanding window of data approach (i.e., 'recursive forecasting'). Each time, we generate forecasts for $h = 1$ up to $h = 17$, which gives forecasts of post-revisions values up to a year ahead, and of the post-revision values of the observation at the forecast origin and of the previous 12 observations.

The resulting 45 sequences of 1 to 17-step ahead forecasts are evaluated against the 2010:1 vintage outturns.

Table 2 records the results for output growth and inflation, measured by the GDP deflator. We provide MSFEs values of the benchmark model, as well as ratios of the MSFEs for the four VAR models to the benchmark. When computing forecasts for the SV-VAR and the SBV-VAR, we assume that we know the future values of $D_s^{T+1|T}, ..., D_s^{T+h|T}$. This means that the forecasts are conditional on the dates of the annual and benchmark revisions being assumed known: at least for the former this seems unobjectionable.

When forecasting output growth, the V-VAR and the RV-VAR both outperform the benchmark in terms of predicting future observations by over 5% on MSFE. But the use of past vintages of data does not improve forecasts of future inflation observations.

The V-VAR improves upon the benchmark (and the RV-VAR) in terms of predicting data revisions to past output growth at all maturities other than the revisions to the first estimate. This suggests that data revisions to output growth are predictable, specially after the second release y_t^{t+2} has been published (the V-VAR versus the RV-VAR). In the case of inflation, the V-VAR beats the benchmark for virtually all maturities. For long maturities (corresponding to observations $T-10$ up to $T-12$), the SBV-VAR is the preferred model, but for all other maturities the V-VAR dominates. Data revisions to GDP inflation are predictable - the V-VAR forecasts register an improvement on MSFE over the 'no-revision' benchmark in excess of 20% for the more recent past ($T-1$ to $T-8$).

The modelling of the periodicity of annual and benchmark revisions improves forecasts of data revisions at short horizons for inflation, but the SV-VAR and the SBV-VAR generally perform worse than the V-VAR and the RV-VAR. The generally disappointing performance of the SV-VAR and the SBV-VAR might be attributable to two reasons.

First, some sort of 'shrinkage' has often been found to improve VAR forecasts (e.g., the Bayesian shrinkage approach of Doan, Litterman and Sims (1984)). Estimating the equation for y_{t-2}^{t+1} of the V-VAR using all observations - those for which y_{t-2}^{t+1} is equal to y_{t-2}^t, as well as those for which y_{t-2}^{t+1} is generated more generally by $\mathbf{y}^t, \mathbf{y}^{t-1} ...,$ ($t+1 \in Q3$) amounts to a form of shrinkage, which may pay dividends in terms of forecast accuracy. Separating out the periods corresponding to annual revisions from those in which no changes are anticipated, as in the S(B)V model, does not appear to be a helpful strategy.

Second, it is generally understood that setting what may be small population coefficients to zero (as in the RV-VAR, relative to the V-VAR or SV-VAR) may produce more accurate forecasts

than the data generating process on squared-error loss criteria (such as MSFE) once an allowance is made for parameter estimation uncertainty (see, e.g., Clements and Hendry (1998) and Giacomini and White (2006)).

In summary, our forecasting results taken together suggest that revisions to past observations on output growth and GDP inflation can be predicted using VAR models, and that future revised output growth can be predicted more accurately than using standard (single-vintage) models. Overall, the V-VAR outperforms the RV-VAR and the SV and SBV-VAR models.

4   Computation of Output Gap in Real Time

There are two reasons why it is inherently difficult to obtain reliable estimates of the output gap in real time (see, e.g., Orphanides and van Norden (2002) and Watson (2007)). The first is the 'one-sided' nature of the data available in real time - real-time estimates of the current value of gaps and trends necessarily have to be made without recourse to future data. However, for data which are autocorrelated, the accurate estimation of gaps and trends requires observations in the future relative to the period of interest. As a consequence, real-time estimates will be markedly less accurate than are possible in historical analyses, when the estimation is based on known future values.[7] The second problem is that gaps and trends are typically defined with reference to 'final' or post-revision data, whereas at time we only have the vintage t values of past data observations. Watson (2007) analyses the first problem, and Garratt *et al.* (2008) also allow for data revisions. Forecast-augmentation is a solution to the one-sidedness problem, whereby forecasts are used to replace unknown future values, and in the context of data subject to revision, Garratt *et al.* (2008) investigate the gains in accuracy that can be achieved by forecasting the revised values of data (as opposed to using a 'single-vintage' AR model to generate the forecasts).

We carry out an exercise similar to that of Garratt *et al.* (2008) to gauge the value of our VAR models in improving the estimates of output gaps and business cycles in real time. We compare the real-time estimates to 'actual' measures of the gap. The latter are calculated using all the historical data from the latest-available vintage, as described below.

Before considering the impact of forecast augmentation, we present results for two one-sided

---

[7]Of course sample end-point problems remain in historical analyses, as the data are one-sided for the estimation of gaps and trends at the beginning and end of the span of historical data.

sets of estimates. The real-time estimate is calculated as:

$$gap_T^{T+1} = filter(y_{1960:1}^{T+1}, ..., y_T^{T+1}) \tag{6}$$

where $filter(a)$ means that we apply a filter to the time series a to extract an unobserved cycle component, and gap_T^{T+1} is set equal to the last of the filtered estimates, which correspond to the y_T observation. Hence the real-time estimate of the gap for observation T uses an estimate of the gap based on vintage-$T + 1$ past values of y_t. The 'pseudo real-time' estimates also use observations up to T, but these are drawn from the 2010:Q1 vintage:

$$gap_T^{2010:1} = filter(y_{1960:1}^{2010:1}, ..., y_T^{2010:1}). \tag{7}$$

The impact of data revision is measured by comparing (7) and (6) - the 'pseudo real-time' and 'real-time' estimates, while the impact of the one-sidedness of the filter is brought out by comparing (7) and the 'true value' (8) defined by:

$$gap_T = filter(y_{1960:1}^{2010:1}, ..., y_{2009:4}^{2010:1}). \tag{8}$$

That is, the gap is computed by applying the filter to the 2010:1-vintage of data covering the whole historical period, 1960:1 to 2009:4, and refers to the observation T.

We consider estimates of the gap for 45 time periods, $T \in \{1995:2, 1995:3,...,2006:2\}$. We use a band-pass filter as Watson (2007) for illustrative purposes. The first plots in Figures 1 and 2 show real-time, pseudo real-time and final estimates of the output gap and the business cycle component.[8] Table 3 presents correlations between the final estimate obtained with historical data and the real-time and pseudo real-time estimates.

The correlation of the 'final' with the 'pseudo real-time' estimates is only 60% for the output gap and 63% for the business cycles, while the use of the real-time vintages reduces these correlation by 5-6 percentage points. Hence we find that the impact of the one-sided nature of the data available in a real-time setting far outweighs data revision effects. However, comparisons of this sort will downplay the importance of data revisions if data revisions can be exploited to generate more accurate forecasts of future observations. That is, the one-sidedness and data revision problems

---

[8]The difference between output gaps and business cycles is that in computing the latter some high frequencies are screened out (for a detailed discussion of the band-pass filter we employ see Watson (2007)). We use the Watson (2007) code available at `www.princeton.edu/~mwatson`.

may be related.

We use forecasts from three VAR models to augment the time series of real output in the period $T+1$ data vintage. Comparing the V-VAR and SBV-VAR enables an assessment of the usefulness of modelling the periodicity of data revisions in terms of the accuracy of derived real-time measures of the output gap. Comparing the results of using the V-VAR and RV-VAR indicates the loss from assuming that revisions after the first are unpredictable (except perhaps for a constant upward or downward shift). All the VAR models have $q = 14$ and $p = 1$, and are estimated with quarterly differences using all the data in the vintages up to and including vintage $T+1$. Forecasts for the level of output are easily recuperated from the growth forecasts.

For each real-time vintage from $T+1$ up to $T+N$ (corresponding to 1995:3 to 2006:3), we generate forecasts of the post-revision values of the next 14 observations, as well as of the revised values of recent past observations.[9] Recall that by post-revision data is meant data which have been revised $q-1$ times, so that y_t^{t+q} is the 'fully-revised' value of y_t. Empirically, we set $q = 14$. So for vintage $T+1$, the gap is calculated as:

$$gap_T^{T+q} = filter(y_{1960:1}^{T+1}, ..., y_{T-q+1}^{T+1}; y_{T-q+2}^{T+2|T+1}, ... y_T^{T+q|T+1}; y_{T+1}^{T+q+1|T+1}, ..., y_{T+q}^{T+q+q|T+1}). \quad (9)$$

The first set of observations in (9) are fully-revised for a given choice of q. The second set are forecasts of the post-revision values of data for which earlier estimates are available. For example, the last of these, $y_T^{T+q|T+1}$, is the forecast of the revised value of y_T, given that we have the first estimate, y_T^{T+1}, in data vintage-$T+1$. The third set are forecasts of the revised values of future observations.

We also consider a variant of (9) that augments the time series of $T+1$ vintage with predictions of data revisions, but omits the forecasts of future observations, to indicate which aspect of (9) is primarily responsible for any improvement.

The results of the forecast-augmented real-time estimates are shown in the second panels of figures 1 and 2. Table 3 indicates that the correlation with the true values for the output gap increases from around 55% to around 71% when the V-VAR is used (equation (9)). This correlation is some 10 percentage points higher than with the pseudo real-time 'one-sided' estimates (7). Using only predictions of data revisions to augment the times series of the $T+1$ vintage yields some

---

[9]We chose to forecast 14 future observations. The last of the 14 future observations for the final vintage (2006:3) is 2009:4 - this is the last data point of the historical sample that is used to estimate the true trend and gap.

modest improvement when the V-VAR is used, highlighting the importance of incorporating the estimates of future revised values when the filter is applied.

The gap and the business cycle measures using the RV-VAR forecasts have a correlation with the true values that are 3 points lower than when the forecasts are computed with the V-VAR specifications. This suggests that imposing the restriction on the model that annual and benchmark revisions are unpredictable leads to less accurate gap estimates. However, the imposition of the periodicity restrictions is again not helpful (V-VAR versus SBV-VAR).

Finally, table 3 reports the results from forecast augmentation using an AR(8), as in Garratt *et al.* (2008). The correlations of the real-time gaps and cycles with the actual improve on the one-sided gaps, but only to the extent that the estimates of the gap are on par with using pseudo real–time data (without forecast augmentation). The performance of the pseudo estimates is actually slightly worse - using forecast augmentation with an AR(8) on the 2010:Q1 vintage data reduces the correlation with the actual relative to just using one-sided estimates. Clearly, there may be better 'single-vintage' models than the AR(8), but this last comparison highlights the benefits of the multiple-vintage VAR approach for real-time gap and cycle estimation.

In summary, we found in section 3 that V-VAR models provide more accurate forecasts of post-revision data than benchmark forecasts which assume no revisions will be made. This is true even when the first-revised data are available, so that subsequent revisions consist only of annual and benchmark revisions. VAR model were shown to improve the accuracy of real-time measures of output gaps and business cycles by both addressing the end-of-sample problem as well as providing better estimates of the lightly revised data available in the recent past (relative to T). We present results for a single filter (albeit a popular choice in the literature), but would expect the advantages to V-VAR forecast-augmentation to hold irrespective of the filter being used (as the same filter is used for the real-time and final estimates).

5   Computing Inflation Trend and Gap in Real Time

Our approach to estimating the trend inflation and the inflation gap in real time is based on the model of trend/cycle in Stock and Watson (2007, 2010), but see also Cogley *et al.* (2010). Stock and Watson (2010) show that the unemployment gap is a useful predictor of the deviations of inflation from trend (i.e., the inflation gap), where the trend is interpreted as a measure of long-run expected inflation. This is shown to hold for a number of measures of US inflation, including the

GDP deflator, the PCE deflator, the core PCE deflator (the PCE deflator excluding energy and food), and the CPI. However, the decomposition of inflation into trend and gap is not computed in real time, based as it is on the last vintage of data (the 2010:Q3 vintage in their case). Their pseudo-real-time estimates may differ from genuine real-time estimates of the trend and cycle, because the GDP, PCE and core PCE deflators are all based on national accounts data and are subject to revision. Croushore (2008) undertakes a careful study of the nature of data revisions to core PCE inflation (the preferred measure of Stock and Watson (2010)), and finds that revisions are predictable. Moreover, they generally raise the inflation rate. We consider $i)$ whether data revisions have an important impact on the measurement of the inflation trend and cycle, and $ii)$ whether forecasting data revisions using V-VAR models can improve the accuracy of real-time estimates of these quantities. We will consider the GDP deflator and the PCE and core PCE deflators. In section 3 we found that revisions to the GDP deflator were predictable, matching the evidence for the (core) PCE of Croushore (2008).

We begin by outlining the trend-gap decomposition. The observed inflation rate is decomposed into a trend and cycle using the Stock and Watson (2007) integrated moving average $(IMA(1,1))$ model allowing for stochastic volatility. The $IMA(1,1)$ is a reduced form representation of the following structural trend-cycle model:

$$
\begin{aligned}
\pi_t &= \tau_t + \eta_t; \ E(\eta_t) = 0, \ var(\eta_t) = \sigma_{\eta,t}^2 \\
\tau_t &= \tau_{t-1} + \varepsilon_t; \ E(\varepsilon_t) = 0, \ var(\varepsilon_t) = \sigma_{\varepsilon,t}^2, \ cov(\eta_t, \varepsilon_t) = 0,
\end{aligned}
$$

where π_t is the one-quarter rate of inflation (at an annualised rate). τ_t is the trend, which is a random walk driven by a conditionally heteroscedastic disturbance, ε_t, and η_t is the cycle, which is also conditionally heteroscedastic. Stock and Watson (2010) interpret the filtered estimate of the trend, $\tau_{t|t}$, as a measure of the long-run expected inflation rate. This filtered estimate is a one-sided measure of the trend, as it is based only on data up to period t. The gap in the annual rate of inflation at period t using the one-sided trend is:

$$
gap_{\pi,t} = \frac{1}{4} \sum_{i=0}^{3} (\pi_{t-i} - \tau_{t-i|t-i}).
$$

where $\frac{1}{4} \sum_{i=0}^{3} \pi_{t-i}$ is the annual rate of observed inflation at period t.

We calculated the one-sided measure of the trend $\tau_{t|t}$ and the gap_t using data from the latest

available vintage (2010:Q1), with an initial observation period of 1960:Q1.[10]

The real-time measures of inflation trend and gap are obtained with each data vintage from 1965:Q4 onwards for the GDP and the PCE deflators, and from 1996:Q1 onwards for the core PCE deflator, with 1960:Q1 as the first observation. We denote the real-time estimate of the inflation trend at time T, using data vintage $T + 1$, by $\tau_{T|T}^{T+1}$. This is computed from the estimation of the trend-cycle model with stochastic volatility using data from the $T + 1$-vintage. The estimation of the model gives $\tau_{t|t}^{T+1}$, where $t \leq T$- so that $\tau_{T|T}^{T+1}$ is the estimate of the trend at the end point.

The real-time inflation gap is:

$$gap_{\pi,T}^{T+1} = \frac{1}{4} \sum_{i=0}^{3} (\pi_{T-i}^{T+1} - \tau_{T-i|T-i}^{T+1}),$$

where, as before, only data from the $T + 1$ vintage is used.

Figures 3 and 4 present one-sided estimates based on the post-revision data, namely $gap_{\pi,t}$, $\tau_{T|T}$, and real-time estimates, $gap_{\pi,T}^{T+1}$, $\tau_{T|T}^{T+1}$, calculated for vintages from 1965:Q3 up to 2006:Q3. As the data are subject to regular revisions up to 13 quarters after the publication of the first release, we do not show estimates for vintages from 2006:Q4 onwards because post-revision data are not yet available for observations from 2006:Q3 onwards using 2010:Q1 as the last available vintage. It is clear that real-time estimates of the trend are noisier than one-sided estimates, and that there are windows of 2-3 years when the real-time measure is either always below or above the one-sided final estimate. The gap estimates suggest that the divergence between real-time and one-sided final measures has decreased after 1985.

Table 4 records the biases of the real-time estimates (assuming the one-sided final are the true values) for a shorter period from 1995:Q2 up to 2006:Q2 (45 observations). This period is chosen to match the out-of-sample period in section 3. The real-time estimates of both the trend and the gap are downward biased. The real-time GDP inflation trend $\tau_{T|T}^{T+1}$ is nearly a quarter of a percentage point less than the final estimate $\tau_{T|T}$. The gap biases are generally smaller, but they are statistically significant for the GDP and core PCE deflators. Figures 5 and 6 present the data for this subsample: both trend and gap real-time estimates are generally lower than the final estimates from 2001:Q4 onwards. For example, if the real-time bias of core PCE inflation is computed from 2001:Q4 onwards, it is nearly $\frac{1}{4}\%$ point for the trend, and roughly one eighth for the cycle. At a time of low inflation rates, 'errors' of this sort of magnitude due to data revisions

---

[10] We used the code made available on Mark Watson's webpage to replicate the results of Stock and Watson (2010). The code estimates the trend $\tau_{t|t}$ using a MCMC algorithm applied to the one-quarter rate of inflation.

are economically relevant. There are episodes when the two measures gave different signals about inflationary pressures in the economy. During the period 2004:Q1-2005:Q2, the real-time measure of the core PCE inflation cycle suggested that the current inflation was near the expected long-run inflation, while the final measure indicated that inflation was nearly a third of a percentage point above.

5.1 Using vintage-based VARs to improve real-time estimates

Given that revisions to inflation measures are predictable, to what extent can the biases in the real-time estimates of trend and gap be reduced by forecasting data revisions? We replace the data in the $T + 1$-vintage that are still subject to the usual rounds of data revisions with forecasts of their post-revision values. We assume the qth estimates are the post-revision values, so for example, with $q = 14$, the last observation π_T^{T+1} would be replaced with the forecast value $\hat{\pi}_T^{T+14|T+1}$. More generally, the last $q-1$ observations of the vintage $T+1$ are not fully revised in this sense, and these observations are replaced with $\hat{\pi}_{T-q+2}^{T+2|T+1}, \hat{\pi}_{T-q+3}^{T+3|T+1}, ..., \hat{\pi}_T^{T+q|T+1}$. The trend is estimated using:

$$\hat{\pi}_t^{T+q} = \tau_t^{T+q} + \eta_t^{T+q}$$

where $\hat{\pi}_t^{T+q} = \pi_t^{T+1}$ for $t = 1, ..., T - q + 1$, and $\hat{\pi}_t^{T+q} = \hat{\pi}_t^{t+q|T+1}$ for $t = T - q + 2, ..., T$. The gap is:

$$gap_{\pi,T}^{T+q} = \frac{1}{4} \sum_{i=0}^{3} (\hat{\pi}_{T-i}^{T+q} - \tau_{T-i|T-i}^{T+q}).$$

Three vintage-based VAR specifications are used to generate forecasts conditional on each data vintage from 1995:Q3 up to 2006:Q3: the V-VAR, RV-VAR and SBV-VAR. Based on the relative forecasting performance of these models for GDP inflation in section 3, we expect the V-VAR data revision forecasts to be the most beneficial. Croushore (2008) finds that data revisions to PCE inflation can be predicted up to the first annual revision. In that case specifications such as the V-VAR and the SBV-VAR might be better than the RV-VAR for PCE inflation. Recall that the RV-VAR assumes that only the first revision can be predicted based on past vintages.

Table 4 presents the biases of the model augmented real-time estimates $\tau_{T|T}^{T+q}$ and $gap_{\pi,T}^{T+q}$ with respect to the final estimates $\tau_{T|T}$ and $gap_{\pi,T}$ for GDP and PCE inflation. We do not present results for core PCE inflation because vintages of data are only available from 1996:Q1, which is too short a sample to estimate the vintage-based VAR models on. Both the V-VAR and the RV-VAR forecasts of revisions reduce the real-time biases in the estimates of trend and cycle. In

the case of GDP inflation, the trend bias becomes statistically insignificant regardless of which of the models is used to predict future data releases.

The second panel of Table 4 presents the ratios of the MSFEs of the real-time estimates augmented by data revision forecasts to the MSFE of the real-time estimates. As before, we take the one-sided final estimates as the true values. Table entries less than one indicate that the model is successful in reducing the measurement error associated with the use of real-time data. The V-VAR and RV-VAR forecasts reduce the MSFE for the trend by 20% for the GDP deflator, and the V-VAR clearly outperforms the RV-VAR for the PCE deflator. When estimating the gap, the gains are smaller but still marked (e.g., around 8% for the RV-VAR for both inflation measures).

The SBV-VAR forecasts are not at all helpful, and worsen the real-time estimates of the trend and gap, confirming the forecasting performance described and discussed in section 3. Figures 5 and 6 show that the improvements in estimating the trend from using the V-VAR model vary across the period.

In summary: Real-time estimates of the trend and cycle of GDP, PCE and core PCE inflation are downward biased. The real-time measures can be improved by using multi-vintage VARs to predict data revisions. The use of the V-VAR model in real time removes 2/3 of the bias in computing trend inflation and reduces the noisiness of the real-time estimate of the gap for the period 1995-2006.

6   Conclusions

We show that real-time estimates of the output gap, and the inflation trend and gap, can be improved by using past data vintages to predict both revisions to past data and post-revision values of future observations. Because our models use only information on the variable in question, we are able to establish that improvements in the measurement of these key quantities are achievable simply by modelling the dynamics of the data releases published by the statistical agency. Hence for real-time policy analysis we conclude that earlier-vintage data usefully supplements the latest-available vintage of data. Information from other macroeconomic variables is also likely to play a role (see e.g., Cunningham *et al.* (2009)), but we do not consider that issue. We regard the vintage-based VAR as a simple way of incorporating past-vintage information, because the unrestricted model can be estimated by OLS and the forecasts can be calculated using standard econometric software packages.

We report results for just one method of extracting estimates of the unobserved quantities of interest in the case of both output and inflation. Our interest is not in investigating the sensitivity of the estimates of these unobserved variables to the method used, but in establishing that the predictability of revisions to past data and of post revision values of future data (in the case of output) suggests that in general it will be possible to improve real-time estimates of output gap and the inflation trend irrespective of the specific filter or model used. Structural models are commonly used for historical analyses, e.g., Ireland (2007) uses a DSGE model to obtain measures of the output gap and a time-varying inflation target rate, but typically use the latest-available data vintage. In terms of the output gap, Orphanides and van Norden (2002) show that the results of historical analyses might be very misleading compared to what would have appeared to have been the case in real time. We show that the same is true of estimates of the inflation trend and gap. The positive aspect of our results is that the real-time estimates of both the output and inflation gaps and trends can be made closer to the historical estimates by the use of V-VAR model forecasts.

Our evaluation of the forecast performance of the different VAR models also yields some additional insights. Revisions to past data are predictable - the VAR models generally improve upon the 'no-change' benchmark forecasts. Moreover, data revisions that occur after the estimate made available two quarters after the observation period are also predictable, implying that the US BEA annual revisions are in part predictable. It is hard to improve upon the unrestricted VAR model by attempting to better approximate the seasonal nature of the release of data revisions, and we conjecture why this might be the case. Finally, there are interesting differences between the nature of revisions to output growth and inflation, in that our VAR models are able to predict revisions to past and future observations on output growth, but only to past data for inflation.

A   Appendix. Estimation of vintage-based VARs

The V–VAR is an unrestricted VAR, so OLS applied equation-by-equation is adequate. Because the elements of the vector of disturbances ε_t are likely to be correlated, OLS applied to any of the restricted specifications (RV-VAR, SV-VAR and SBV-VAR) is not efficient, and instead estimation of these models is carried out by the seemingly unrelated regression estimator (SURE), as this is equivalent to maximum likelihood (e.g., Hamilton (1994, p. 317)).

In the case of the periodic VARs (SV-VAR and SBV-VAR), we proceed as follows. Define the $q \times k$ matrix of explanatory variables, when $p = 1$, by:

$$\mathbf{x}^t = \begin{bmatrix} (1, \mathbf{y}^{t\prime}) & \mathbf{0}_{1 \times (q+1)} & \mathbf{0}_{1 \times (q+1)} & \cdots & \mathbf{0}_{1 \times (q+1)} \\ \mathbf{0}_{1 \times (q+1)} & (1, \mathbf{y}^{t\prime}) & \mathbf{0}_{1 \times (q+1)} & \cdots & \mathbf{0}_{1 \times (q+1)} \\ \mathbf{0}_{1 \times (q+1)} & \mathbf{0}_{1 \times (q+1)} & (D_s^{t+1}, \mathbf{y}^{t\prime} \times D_s^{t+1}) & & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times (q+1)} & \mathbf{0}_{1 \times (q+1)} & \mathbf{0}_{1 \times (q+1)} & \cdots & (D_s^{t+1}, \mathbf{y}^{t\prime} \times D_s^{t+1}) \end{bmatrix}$$

where $k = q(q + 1)$. The model is then given by:

$$\mathbf{y}^{t+1} = \mathbf{A}\mathbf{y}^t \times (1 - D_s^{t+1}) + \mathbf{x}^t \boldsymbol{\beta} + \mathbf{v}^{t+1}$$

where:

$$A_{q \times q} = \begin{bmatrix} 0_{2 \times 1} & 0_{2 \times q-2} & 0_{2 \times 1} \\ 0_{q-2 \times 1} & I_{q-2 \times q-2} & 0_{q-2 \times 1} \end{bmatrix}.$$

In order to estimate such a model, we define $\mathbf{z}_t = (\mathbf{y}^{t+1} - \mathbf{A}\mathbf{y}^t * (1 - D_s^{t+1}))$, and apply the SURE estimator to the system:

$$\mathbf{z}_t = \mathbf{x}_t \boldsymbol{\beta} + \mathbf{v}_t.$$

References

Clark, T. E. (2010). Real-time density forecasting from BVARs with stochastic volatility. *Federal Reserve of Kansas City, mimeo.*

Clements, M. P., and Galvão, A. B. (2010). Real-time forecasting of inflation and output growth in the presence of data revisions. Warwick economics research paper, no. 953, Department of Economics, University of Warwick.

Clements, M. P., and Galvão, A. B. (2011). Forecasting with vector autoregressive models of data vintages: US output growth and inflation. Discussion paper, Department of Economics, University of Warwick.

Clements, M. P., and Hendry, D. F. (1998). *Forecasting Economic Time Series.* Cambridge: Cambridge University Press. The Marshall Lectures on Economic Forecasting.

Clements, M. P., and Hendry, D. F. (2006). Forecasting with breaks. In Elliott, G., Granger, C., and Timmermann, A. (eds.), *Handbook of Economic Forecasting, Volume 1. Handbook of Economics 24*, pp. 605–657: Elsevier, Horth-Holland.

Cogley, T., Primiceri, G. E., and Sargent, T. J. (2010). Inflation-Gap Persistence in the US. *American Economic Journal: Macroeconomics*, **2:1**, 43–69.

Croushore, D. (2008). Revisions to PCE inflation measures: Implications for monetary policy. Working paper no. 08-8, Research Department, Federal Reserve Bank of Philadelphia.

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

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.*

Doan, T., Litterman, R., and Sims, C. A. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometrica*, **3**, 1–100.

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.

Giacomini, R., and White, H. (2006). Tests of conditional predictive ability. *Econometrica*, **74**, 1545 – 1578.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton: Princeton University Press.

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.

Ireland, P. N. (2007). Changes in the Federal Reserve's inflation target: Causes and consequences. *Journal of Money, Credit and Banking*, **39**, 1851–1882.

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.

Kozicki, S., and Tinsley, P. A. (2005). Permanent and transitory policy shocks in an empirical macro model with asymmetric information. *Journal of Economic Dynamics and Control*, **29**, 1985–2015.

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

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

Orphanides, A., and van Norden, S. (2002). The unreliability of output gap estimates in real time. *The Review of Economics and Statistics*, **84**, 569–583.

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.

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

Stock, J. H., and Watson, M. W. (2007). Why has U.S. Inflation Become Harder to Forecast?. *Journal of Money, Credit and Banking*, **Supplement to Vol. 39**, 3–33.

Stock, J. H., and Watson, M. W. (2010). Modelling Inflation after the Crisis. *NBER Working Paper Series*, **16488**.

Watson, M. W. (2007). How accurate are real-time estimates of output trends and gaps?. *Federal Reserve Bank of Richmond Economic Quarterly*, **93**, 143–161.

Table 1. In-sample fit of the models

	$ln\left|\hat{\Sigma}_{\varepsilon}\right|$	n. param.	H$_0$: RV-VAR
	Output growth		
RV-VAR	-23.92	42	_
V-VAR	-25.26	210	237.29 [0.000]
SV-VAR	-29.90	210	1059.2 [0.000]
SBV-VAR	-31.76	210	1387.5 [0.000]
	Inflation		
RV-VAR	-38.72	42	_
V-VAR	-40.11	210	245.40 [0.000]
SV-VAR	-44.56	210	1033.1 [0.000]
SBV-VAR	-46.89	210	1446.0 [0.000]

Notes: The column labelled `H$_0$: RV-VAR' gives the LR statistics for the null that the true model is the RV-VAR against each model, while the values in brackets are the *p*-values of the null. All models have *p=1*, *q=14*, and are estimated on the data vintages from 1965:Q4 up to 2010:Q1.

Table 2. Comparing forecasts of post-revision output growth and inflation using data from the 2010:Q1 vintage to compute forecast errors.

h	obs.	Output Growth					Inflation (GDP deflator)				
		Bench.	V-VAR	RV-VAR	SV-VAR	SBV-VAR	Bench.	V-VAR	RV-VAR	SV-VAR	SBV-VAR
		Forecasting Post-Revision Future Observations									
17	T+4	4.654	0.931	0.923	1.014	1.096	1.024	0.960	1.037	1.084	1.036
16	T+3	4.565	0.923	0.919	0.993	1.076	0.843	1.051	1.124	1.206	1.127
15	T+2	4.448	0.936	0.943	1.029	1.091	0.818	1.084	1.089	1.196	1.158
14	T+1	4.262	0.939	0.958	1.101	1.082	0.625	0.994	1.002	1.225	1.138
	Sum	17.927	0.932	0.935	1.033	1.086	3.310	1.020	1.065	1.170	1.108
		Forecasting Post-Revision Releases (Data Revisions)									
13	T	2.602	1.064	1.069	1.409	1.284	0.377	0.768	0.835	1.170	1.024
12	T-1	2.136	0.931	1.022	1.432	1.218	0.297	0.816	0.930	1.482	1.361
11	T-2	2.153	0.929	1.026	1.352	1.167	0.323	0.766	0.904	1.246	1.140
10	T-3	2.098	0.959	1.018	1.475	1.221	0.345	0.759	0.927	1.354	0.994
9	T-4	2.243	0.896	1.005	1.311	1.122	0.314	0.806	0.968	1.392	1.008
8	T-5	1.664	0.805	0.974	1.229	1.078	0.317	0.666	1.019	1.319	0.845
7	T-6	1.641	0.868	0.990	1.129	0.972	0.288	0.738	1.020	1.384	0.867
6	T-7	1.504	0.901	0.988	1.231	1.030	0.252	0.726	1.022	1.230	0.779
5	T-8	1.593	0.897	0.972	1.158	0.986	0.165	0.767	1.016	1.163	0.950
4	T-9	1.394	0.894	0.986	0.957	0.925	0.121	0.846	1.013	1.168	1.117
3	T-10	1.229	0.877	0.995	0.876	0.920	0.111	1.021	1.004	1.303	0.973
2	T-11	1.183	0.876	0.980	0.866	0.955	0.116	0.973	1.000	1.108	0.857
1	T-12	0.931	0.946	0.982	0.984	1.009	0.132	1.020	0.986	1.127	0.816
	Sum	22.370	0.919	1.006	1.234	1.097	3.157	0.789	0.959	1.289	0.994

Notes. The models are estimated with increasing windows of data, beginning with data vintages from 1965:4 to 1995:3, and ending with the vintages 1965:4 to 2006:3. The entries in the first column are the MSFEs of the benchmark forecasts, and the remaining columns are ratios of MSFEs relative to the first column entries. The benchmark forecasts are generated from AR models (an AR(1) for output growth, and an (AR(4) for inflation) for future observations (*T+1* to *T+4*) , and a random walk `no-change' predictor for past data (*T* down to *T-12*).

In terms of the main text notation, the post-revision forecasts are $\hat{y}_{T+1+h-14}^{T+1+h|T+1}$, with *q=14*. So when *h=1* we have a forecast of the post-revision value at y_{T-12} , and when *h=17*, of the post-revision value of y_{T+4} (relative to the *T+1*-data vintage).

Table 3. Estimating output gaps and business cycles in real time - Correlations of real-time estimates with historical estimates.

	Output gap (filter includes periods: 2-32)		Business cycles (filter includes periods: 6-32)	
		w/AR(8)		w/AR(8)
Real-time	0.55	0.59	0.57	0.63
Pseudo	0.60	0.57	0.63	0.59
Aug. with	$\hat{y}_{T-13}^{T+2|T+1}, ..., \hat{y}_T^{T+14|T+1}$	$\hat{y}_{T-13}^{T+2|T+1}, ..., \hat{y}_T^{T+14|T+1}$ $\hat{y}_{T+1}^{T+15|T+1}, ..., \hat{y}_{T+14}^{T+28|T+1}$	$\hat{y}_{T-13}^{T+2|T+1}, ..., \hat{y}_T^{T+14|T+1}$	$\hat{y}_{T-13}^{T+2|T+1}, ..., \hat{y}_T^{T+14|T+1}$ $\hat{y}_{T+1}^{T+15|T+1}, ..., \hat{y}_{T+14}^{T+28|T+1}$
V-VAR	0.58	0.71	0.60	0.71
RV-VAR	0.56	0.68	0.58	0.68
SBV-VAR	0.55	0.68	0.58	0.69

Notes. The true/historical estimates use the 2010:1 data vintage. The gap and cycle estimates are computed for the periods (vintages) 1995:Q2(Q3)-2006:Q2(Q3), using data observations from 1960:1 onwards. w/AR(8) means that the (pseudo) real-time series of output has being augmented with forecasts using an AR(8) model, before applying the band-pass filter. The output series is augmented by 14 future observations in all cases that augmentation is implemented. Hence at 2006:Q2, for example, the series is predicted up to 2009:Q4 (which is the last observation used for the historical estimate). The code on the band-pass filter is as used by Watson (2007), with exactly the same cut-offs to obtain gap and business cycles.

Table 4. Measurement of the inflation trend and gap in real-time: the role of the VAR forecasts.

Table 4.1. Bias.

	Trend				Gap			
	Real-time	Aug. with: $\hat{y}_{T-13}^{T+2|T+1}, \dots, \hat{y}_T^{T+14|T+1}$			Real-time	Aug. with: $\hat{y}_{T-13}^{T+2|T+1}, \dots, \hat{y}_T^{T+14|T+1}$		
		V-VAR	RV-VAR	SBV-VAR		V-VAR	RV-VAR	SBV-VAR
GDP deflator	0.223 **[.016]**	0.071 [.449]	0.093 [.318]	0.006 [.952]	0.057 **[.004]**	0.032 [.089]	0.029 [.141]	0.029 [.335]
PCE deflator	0.120 **[.018]**	0.065 [.202]	0.005 [.920]	-0.132 **[.073]**	0.037 [.266]	0.012 [.711]	0.015 [.642]	0.013 [.727]
PCE core	0.144 **[.006]**				0.078 **[.009]**			

Table 4.2. Ratio of MSFEs of the forecast-augmented real-time estimates to standard real-time estimates.

	Trend			Gap		
	V-VAR	RV-VAR	SBV-VAR	V-VAR	RV-VAR	SBV-VAR
GDP deflator	0.797	0.806	1.186	0.836	0.920	2.077
PCE deflator	0.710	0.891	1.374	1.006	0.917	1.287

Notes. For both the bias and MSFE calculations the 2010:1 data vintage are the `actual values' for the calculation of forecast errors. The trend and gap estimates are for the period (vintages) 1995:2(3)-2006:2(3), using data from 1960:1. The model forecasts are used to provide post-revision estimates of the last *q-1* observations at each forecast origin which are subject to revision. In table 4.1 [.] denotes *p*-values of the test of zero bias (with Newey-West standard errors). The MSFE and bias calculations refer to the trend and cycle expressed at annual rates.

Figure 1: Output Gap with vintages from 1995:Q3 (T+1) up to 2006:Q3 (T+N).

1.1. Final: gap_T , **pseudo real-time:** $gap_T^{2010:1}$, **real-time:** gap_T^{T+1}

1.2 Final: gap_T ; **model augmented real-time estimates with: V-VAR, RV-VAR and SBV-VAR.**
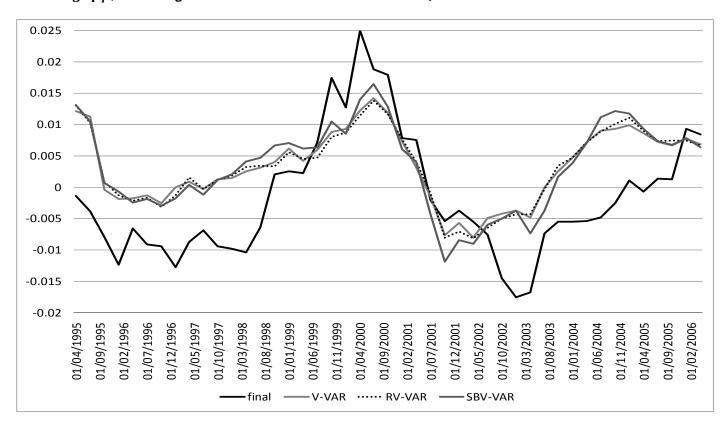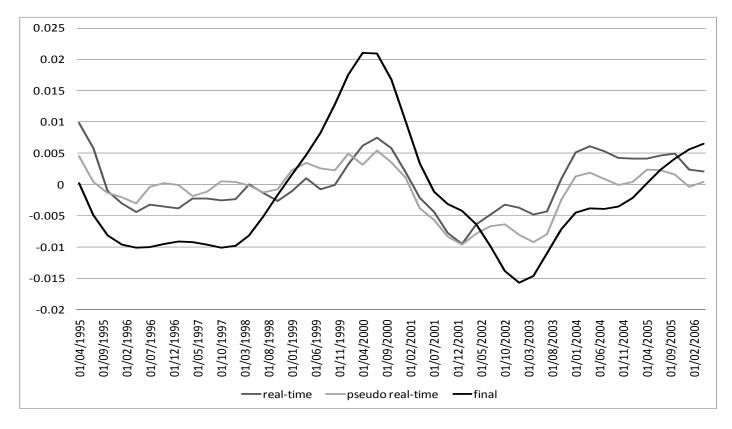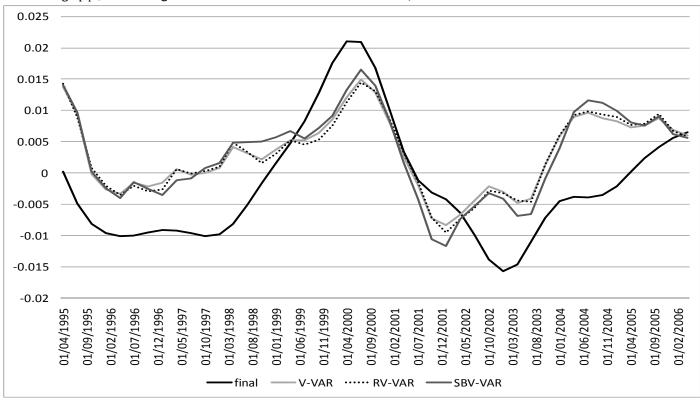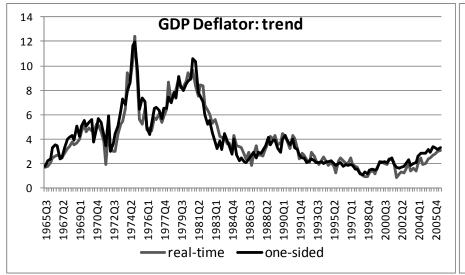
Figure 2: Business Cycles with vintages from 1995:Q3 (T+1) up to 2006:Q3 (T+N).

2.1. Final: gap_T , pseudo real-time: $gap_T^{2010:1}$, real-time: gap_T^{T+1}

2.2. Final: gap_T ; model augmented real-time estimates with: V-VAR, RV-VAR and SBV-VAR.

Figure 3: Real-time $\tau_{T|T}^{T+1}$ and final one-sided $\tau_{T|T}$ estimates of trend inflation.

GDP Deflator: trend

PCE Deflator: trend

PCE core: trend

Figure 4: Real time $gap_{\pi,T}^{T+1}$ and final one-sided $gap_{\pi,T}$ estimates of inflation gap.
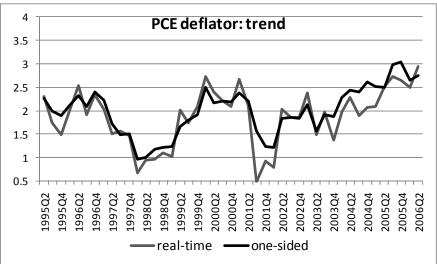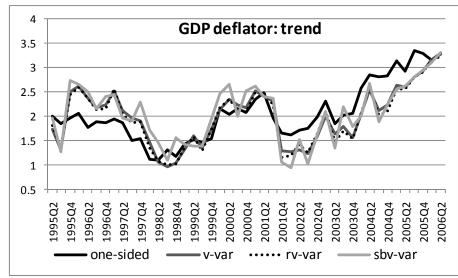
Figure 5: Real time and final one-sided estimates of trend inflation; model augmented real-time estimates of inflation trend: V-VAR, RV-VAR and SBV-VAR.
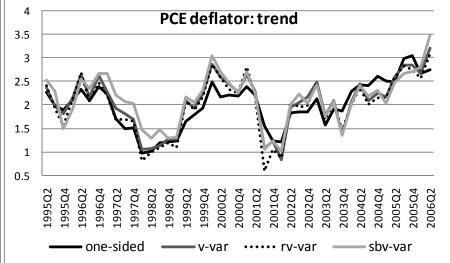
Figure 6: Real time and final one-sided estimates of inflation gap; model augmented real-time estimates of inflation gap: V-VAR, RV-VAR and SBV-VAR.