First Announcements and Real Economic Activity

Michael P. Clements and Ana Beatriz Galvão

No 885

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

First Announcements and Real Economic Activity

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

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

December 18, 2008

Abstract

The recent literature suggests that first announcements of real output growth in the US have predictive power for the future course of the economy. We show that this need not point to a behavioural relationship, whereby agents respond to the announcement, but may instead simply be a by-product of the data revision process. Initial estimates are subsequently subject to a number of rounds of revisions: the nature of these revisions is shown to be key in determining any apparent relationship between first announcements and the future course of the economy.

Keywords. Announcements, real activity, data measurement, revisions.
JEL code: C52

*Corresponding author: Ana Beatriz Galvão, Queen Mary University of London, Department of Economics, Mile End Road, E1 4NS, London, UK, Phone: ++44-2074-78828255, a.ferreira@qmul.ac.uk

1 Introduction

The key issue we investigate in this paper is whether agents' beliefs or perceptions about the current state of the economy affect the future evolution of the economy. In a recent paper, Rodriquez Mora and Schulstad (2007, p. 1934) argue that:

> The beliefs of agents (expressed in the announcement, their available information) determine the future path of the economy much more than the true events that the announcement measures.

The assertion is based on the empirical evidence that last quarter's actual rate of output growth has no predictive power for the present quarter's actual rate of output growth once we take into account the first announcement of the rate of output growth in the last quarter. They relate their results to those of Oh and Waldman (1990, 2005), who report that announcements of the leading economic indicator affect future activity (as measured by industrial production), and that these announcements are an important source of expectational shocks. Specifically, Oh and Waldman (2005) show that errors in the initial announcements of the leading indicator affect survey expectations of future activity. Referring to the literature on strategic complementarity, they state that [p.75]:

> if all agents suddenly revised upwards their beliefs concerning the production plans of other agents - even if there were no change in any real variable - the result would be a type of self-fulfilling increase in the future level of production

which would be consistent with their earlier findings that false announcements affect the future evolution of the economy. As they note, the importance of expectational errors for business cycle fluctuations dates back to an earlier literature associated with such luminaries as Keynes and Pigou, as well as the more recent literature on sunspot equilibria and strategic complementarities (see Oh and Waldman (2005) for references).

In this paper we consider an alternative explanation of the finding by Rodriquez Mora and Schulstad (2007) which stresses issues to do with measurement, rather than behavioural explanations which suppose agents respond to announcements irrespective of the true state of the economy. A reason for considering measurement issues is simply that revisions to estimates of the true state of the economy are ongoing, so that the 'true' GDP growth rate in any period will not be determined until a number of years later, and will typically depend on a whole raft of revisions, as discussed below. There is the very real possibility that the nature of the revisions process may of itself generate the findings that Rodriquez Mora and Schulstad (2007) report.

There has been much interest in the recent literature in the nature of data revisions and the revisions process, and whether early estimates are efficient forecasts of later, revised estimates, or are simply noisy estimates of 'final' values (see, e.g., Mankiw and Shapiro (1986), Aruoba (2008), as well as the review in Jacobs and van Norden (2006)). We will consider how the findings of Rodriquez Mora and Schulstad (2007) fit within this literature on 'measurement'.[1]

As well as issues to do with the nature of the revisions process, the question of whether data vintages matter has also been addressed - do key macroeconomic results or relationships established for one particular vintage of data remain relevant for other vintages of data. For example, Croushore and Stark (2003) examine three major studies in macroeconomics, and find that of these three the results of the seminal paper by Hall (1978) on the rational expectations permanent income hypothesis appear to be dependent on the particular data vintage studied. Runkle (1998), Orphanides (2001) and Orphanides and van Norden (2005) consider the effects of data revisions on the conduct of monetary policy and the calculation of output gaps. There has also been much interest in data vintages and forecasting, as the use of final-revised data may give a misleading impression relative to the use of data available at the time in pseudo real-time forecasting exercises (see, for example, Diebold and Rudebusch (1991), Faust, Rogers and Wright (2003), and the recent review by Croushore (2006)). Recently, a number of authors have considered how to specify forecasting models when there are various data vintage estimates of the same observation (see, e.g., Koenig, Dolmas and Piger (2003), Clements and Galvão (2008b) and Clements and Galvão (2008a)).

Relative to this burgeoning literature on measurement in the presence of multiple data vintages, consideration of the impact of early vintage estimates on the final value of real output in subsequent quarters gives rise to a different line of analysis. Our interest is not in whether the same macro relationship holds on different vintages of data, but in a putative relationship that draws on different data vintages, namely first-release data, and final release data. In this paper we consider whether the recent literature on modelling and testing data revisions is compatible with the findings of Rodriquez Mora and Schulstad (2007).

Our main contribution is to show that the dependence of real output growth on prior first announcements can be viewed as a by-product of the data revision process, rather than revealing agents responding to first announcements. We use two different models of the dynamics of the revision process to investigate the relationship between real output growth and first announcements, and the dependence of this relationship on the revisions process. The first model supposes that the true state of the economy is a latent process that may never be observed, and is based on the state-space model of Jacobs and van Norden (2006) and allows for correlated revisions or 'spillovers'. The

[1] Rodriquez Mora and Schulstad (2007, p. 1926) talk of 'noise' in the first announcement, but it appears that they are not using this term in the Mankiw and Shapiro (1986) sense, as discussed below. They do not test for noise, although their regression (2) of the final value on the first announcement is a test of news (this is formally equivalent to the more familiar form of regressing the revision on the first announcement), and they reject the news hypothesis.

second is a vector autoregression that models the various data vintages directly. Both approaches allow the identification and testing of whether data revisions incorporate news to the measurement of output growth. Of interest are two related questions: whether these frameworks are capable of generating the finding that first announcements affect the future course of the economy; and assuming that they do, whether the predictive power of first announcements holds for empirically relevant regions of the parameter space.

The next section describes the main findings of Rodriquez Mora and Schulstad (2007), and reproduces their estimates on our data set. Section 3 analyses their findings within a modelling framework that characterises data revisions as news or noise processes. Section 4 investigates the impact of announcements within the vector autoregressive framework which has become the most popular modelling framework for empirical research since its inception by Sims (1980). Finally, section 5 offers some concluding remarks.

2 Perceptions and the economy

Rodriquez Mora and Schulstad (2007) argue that 'perceptions affect the economy' based on the empirical evidence they report that shows that last quarter's actual rate of output growth has no predictive power for the present quarter's actual rate of output growth once we take into account the first announcement of the rate of output growth in the last quarter. By 'actual' is meant the estimate published a number of years later, which is assumed to reveal the truth.

Letting y_t denote the final vintage value of real output growth (GNP/GDP) in quarter t, the regression:

$$y_t = \delta_0 + \delta_1 y_{t-1} + \varepsilon_t \tag{1}$$

over the sample period 1967 to 1991, is found to yield a statistically significant estimate $\hat{\delta}_1$ of around 0.32, giving rise to the standard finding that post WWII US output growth can be approximated by a first-order autoregression. But if the first announcement of growth in $t-1$ is included as an additional explanatory variable, Rodriquez Mora and Schulstad (2007) find that the estimate of the coefficient on y_{t-1} is no longer statistically significantly different from zero. Letting y_{t-1}^t denote the first announcement of the value of output in period $t-1$ (available one period later- denoted by the 't' superscript), then the regression model becomes:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-1}^t + v_t \tag{2}$$

and we fail to reject the null that then $\alpha_1 = 0$, but reject the null that $\alpha_2 = 0$, with $\hat{\alpha}_2$ equal to 0.36. The first announcement has a statistically significant impact on the final value, whilst last period's final value has no predictive power once the first announcement of the value of output in that period is included.

We begin by replicating these findings. Rodriquez Mora and Schulstad (2007) use the 'advance' or first announcement (see their footnote 4), made available soon after the end of the quarter. Our first-release series is taken from the monthly vintages of real output growth made available by the Real-Time Data Set for Macroeconomists (www.philadelphiafed.org, see Croushore and Stark (2001)), and we use the vintage from the second month in the quarter after the quarter of interest (which corresponds to the advance report of the Bureau of Economic Analysis, BEA). Here and throughout we use a number of different definitions of final data because of the ambiguity in this notion. We take the final data to be the value after three years of revisions, after six years of revisions, and the value recorded in the latest available vintage, 2008Q2. We denote these by $\{y_t^{t+12}\}$, $\{y_t^{t+24}\}$ and $\left\{y_t^{2008Q2}\right\}$, respectively, so that y_t^{t+12} is the estimate of the value of y in period t from the data vintage of $t+12$ (12 quarters after the observational period), and so on.

Rodriquez Mora and Schulstad (2007) use GNP data from 1967 to 1991. We will consider a number of different sample periods, the longest of which will be 1965Q3 to 2001Q4. By ending the estimation sample in 2001Q4, all the observations $\left\{y_t^{2008Q2}\right\}$ will have undergone a minimum of seven years of revisions. But the earliest observations (those for the 1960's) will have been subject to nearly forty years of revisions. Every five or ten years data are typically subject to benchmark revisions, reflecting methodological changes in measurement or collection procedures (including base year changes), in addition to the 'regular revisions' process which normally runs for three years.[2] This motivates the use of the $\{y_t^{t+12}\}$ as a partial control for benchmark revisions. The use of $\{y_t^{t+12}\}$ in place of $\left\{y_t^{2008Q2}\right\}$ in (2) should mitigate the impact of benchmark revisions on the relationship between the 'final' and early release. Corradi, Fernandez and Swanson (2007) and van Dijk, Franses and Ravazzolo (2007) both warn of the consequences of benchmark revisions on the relationship between first and final release estimates. The use of $\{y_t^{t+24}\}$, the vintage available six years after the first estimate of each observation of the output growth, as a measure of the final series is motivated in section 4.

Hence, our choice of data vintage and estimation sample should ensure that the main difference between $\{y_t^{t+12}\}$ and both $\{y_t^{t+24}\}$ and $\left\{y_t^{2008Q2}\right\}$ as measures of the final values is in terms of the

[2] Data published by the BEA are normally revised up to three years after they are first released. In July of each year there are revisions to the National Accounts data for the first quarter of the current year, as well as all the quarters of the previous three years. 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. Although we only consider growth rates, the use of a fixed-weighting method prior to the introduction of chain-weighting in 1996 means that the growth rates of real variables changed, especially observations not close to the base year.

effects of benchmark revisions rather than the number of standard revisions they have been subject to.

When $\left\{y_t^{2008Q2}\right\}$ is taken as the series of final values, table 1 reproduces the finding that the lagged actual series is insignificant when the lagged first-release series is included, pointing to the importance of first announcements as opposed to the events which they measure. The size of the coefficient on the lagged first-released data is also larger than that on lagged final-data when instead the final series is measured by $\left\{y_t^{t+24}\right\}$. Clearly there is a degree of collinearity between the explanatory variables that affects the estimated coefficients' standard errors, whereby even with an R^2 as large as 16% the null hypotheses of the individual insignificance of both regressors cannot be rejected. Because of multicollinearity, we will consider magnitudes of effects and not just statistical significance, especially when smaller sample sizes are being used. The regressions involving $\left\{y_t^{2008Q2}\right\}$ and $\left\{y_t^{t+24}\right\}$ will be subject to the effects of benchmark revisions. When we estimate the regression (2) using $\left\{y_t^{t+12}\right\}$ for the final data, the dependence of output growth on the 'first announcement' no longer holds. Using the $\left\{y_t^{t+12}\right\}$ series partially controls for benchmark revisions and indicates that lagged first release output growth is not statistically significant, whereas lagged actual is borderline significant at the 10% level.

Table 1 also presents results using the same period as Rodriquez Mora and Schulstad (2007), that is, 1967-1991. The broad pattern of results is confirmed, subject to the caveats about statistical significance. Finally, we report results for the period 1974-2001, as this matches the sample period used in the next two sections. Importantly, we find that the lagged announcements is the key explanatory variable when $\left\{y_t^{2008Q2}\right\}$ and $\left\{y_t^{t+24}\right\}$ are used as final data, and the lagged actual is more important when $\left\{y_t^{t+12}\right\}$ are the final data.

Finally, the last column records the results of regressing the $t + 24$ 'final values' on the $t + 12$ vintage 'first announcements'. A telling finding is that the lagged value of the final series is insignificant, and the 'earlier' announcement is either statistically significant or has an effect four or five times as large, depending on the sample period. Because the earlier announcement is not available until 3 years after the observation to which it refers, the significance of this term in the regression cannot be because agents alter their behaviour in response to it. Instead we suggest that the significance of the earlier announcement term is due to the data revision process. In the remainder of the paper we consider whether the data revision process is also able to explain the significance of the first announcement in the first three columns of table 1.

In the next two sections we analyse whether these findings are consistent with two standard approaches to modelling data revisions process: the state-space model in section 3 and the vector autoregressive model in section 4.

3 A state-space model of data vintages

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 y_t, as well as (in the general case) news and noise components, v_t^{t+s} and ε_t^{t+s}:

$$y_t^{t+s} = y_t + v_t^{t+s} + \varepsilon_t^{t+s}.$$

Data revisions are news when initially released data are optimal forecasts of later data, so news revisions are not correlated with the new released data, that is,

$$Cov\left(v_t^{t+s}, y_t^{t+s}\right) = 0.$$

Data revisions are noise when each new release of the data is equal to the true value y_t plus noise, so that noise revisions are not correlated with the truth:

$$Cov\left(\varepsilon_t^{t+s}, y_t\right) = 0.$$

We adopt the framework of Jacobs and van Norden (2006) 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 the vector of noise revisions $\boldsymbol{\varepsilon}_t = \left(\varepsilon_t^{t+1}, \ldots, \varepsilon_t^{t+l}\right)'$ and the one of news revisions $\mathbf{v}_t = \left(v_t^{t+1}, \ldots, v_t^{t+l}\right)'$, so that $\mathbf{y}_t = y_t + \mathbf{v}_t + \boldsymbol{\varepsilon}_t$. One way of defining a revisions process with the required characteristics is to assume a process for y_t, for example, an AR(1) with iid disturbances η_{1t}, plus a sum of iid disturbances η_{2t}:

$$y_t = \rho_1 y_{t-1} + R_1 \eta_{1t} + \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i}. \tag{3}$$

and then specify $\mathbf{v}_t$ and $\boldsymbol{\varepsilon}_t$ as:

$$\overset{\mathbf{y}_t}{\begin{bmatrix} y_t^{t+1} \\ y_t^{t+2} \\ \\ \\ y_t^{t+l} \end{bmatrix}} = y_t - \overbrace{\begin{bmatrix} \sigma_{v_1} & \sigma_{v_2} & \cdots & \sigma_{v_l} \\ & \sigma_{v_2} & & \\ & & \ddots & \\ & & & \\ 0 & & & \sigma_{v_l} \end{bmatrix} \begin{bmatrix} \eta_{2t,1} \\ \eta_{2t,2} \\ \\ \\ \eta_{2t,l} \end{bmatrix}}^{\{\mathbf{v}_t\}} + \overset{\{\boldsymbol{\varepsilon}_t\}}{\begin{bmatrix} \sigma_{\varepsilon_1}\eta_{3t,1} \\ \sigma_{\varepsilon_2}\eta_{3t,2} \\ \\ \\ \sigma_{\varepsilon_l}\eta_{3t,l} \end{bmatrix}}, \tag{4}$$

where $\boldsymbol{\eta}_t = [\eta_{1t}, \boldsymbol{\eta}_{2t}', \boldsymbol{\eta}_{3t}']$ is iid, $E(\boldsymbol{\eta}_t) = 0$, with $E(\boldsymbol{\eta}_t\boldsymbol{\eta}_t') = I$. Thus $\sigma_{v_1}, ..., \sigma_{v_l}$ are standard deviations of $\eta_{2t,1}, ..., \eta_{2t,l}$ processes, $\sigma_{\varepsilon_1}, ..., \sigma_{\varepsilon_l}$ are standard deviations of $\eta_{3t,1}, ..., \eta_{3t,l}$ processes, and $R_1 = \sigma_{\eta_1}$ is the standard deviation of the disturbances of the underlying AR(1) process for the true values. Therefore, the first estimate of y_t, y_t^{t+1}, estimates y_t with noise $(\sigma_{\varepsilon_1}\eta_{3t,1})$ and a news term consisting of l separate components $(-\sum_{i=1}^{l} \sigma_{v_i}\eta_{2t,i})$. Later estimates are also characterised by noise, but fewer news components and therefore provide more accurate estimates of y_t. If $\sigma_{v_l} = 0$ and $\sigma_{\varepsilon_l} = 0$ the l-vintage value is the true value, $y_t^{t+l} = y_t$.

Combining equation (4) with the statistical process for y_t, equation (3), the specification implies that the set of revisions, $\mathbf{y}_t - \mathbf{i}y_t = \mathbf{v}_t + \boldsymbol{\varepsilon}_t$ ($\mathbf{i}$ an $l \times 1$ vector of ones) is uncorrelated with y_t when there is no news ($\mathbf{v}_t = 0$), i.e., $E(\boldsymbol{\varepsilon}_t y_t) = 0$ by the assumption that $E(\eta_t \eta_t') = I$. As a consequence, when revisions are **pure noise**:

$$\mathbf{y}_t - \mathbf{i}y_t = \boldsymbol{\varepsilon}_t, \text{ so that } E(\boldsymbol{\varepsilon}_t y_t) = 0.$$

Combining equations (4) and (3) also implies that the revisions are uncorrelated with $\mathbf{y}_t$ when there is no noise ($\boldsymbol{\varepsilon}_t = 0$), because then: $y_t^{t+s} = y_t + \nu_t^{t+s} = \rho_1 y_{t-1} + R_1 \eta_{1t} + \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i} - \sum_{i=s}^{l} \sigma_{v_i} \eta_{2t,i} = \rho_1 y_{t-1} + R_1 \eta_{1t} + \sum_{i=1}^{s-1} \sigma_{v_i} \eta_{2t,i}$. Hence $E\left(v_t^{t+s} y_t^{t+s}\right) = 0$ since $\nu_t^s = -\sum_{i=s}^{l} \sigma_{v_i} \eta_{2t,i}$, for $s = 1, \ldots, l$. As a consequence, when revisions are **pure news**:

$$y_t^{t+s} = \rho_1 y_{t-1} + R_1 \eta_{1t} + \sum_{i=1}^{s-1} \sigma_{v_i} \eta_{2t,i}, \text{ so that } E\left(\nu_t^{t+s} y_t^{t+s}\right) = 0.$$

The model can be cast in state-space form (SSF), where the transition equations are described by:

$$\boldsymbol{\beta}_{t+1} = \mathbf{T}\boldsymbol{\beta}_t + \mathbf{R}\eta_{t+1}$$

$$\boldsymbol{\beta}_{t+1} = \begin{bmatrix} y_t \\ y_{t-1} \\ \mathbf{v}_{t+1} \\ \boldsymbol{\varepsilon}_{t+1} \end{bmatrix}, \ \mathbf{T} = \begin{bmatrix} \rho_1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{T}_3 & 0 \\ 0 & 0 & 0 & \mathbf{T}_4 \end{bmatrix}, \ \mathbf{R} = \begin{bmatrix} R_1 & \mathbf{R}_3 & 0 \\ 0 & 0 & 0 \\ 0 & -\mathbf{U}_1.diag(\mathbf{R}_3) & 0 \\ 0 & 0 & \mathbf{R}_4 \end{bmatrix}, \ \eta_t = \begin{bmatrix} \eta_{1t} \\ \boldsymbol{\eta}_{2t} \\ \boldsymbol{\eta}_{3t} \end{bmatrix} \tag{5}$$

Here, $\mathbf{R}_3 = [\sigma_{v_1} \ldots \sigma_{vl}]$, $\mathbf{U}_1$ is an upper-triangular matrix of ones, and $\mathbf{R}_4 = diag\,(\sigma_{\varepsilon_1}, \ldots, \sigma_{\varepsilon_l})$.The measurement equation is:

$$\mathbf{y}_t = y_t + I_l \mathbf{v}_t + I_l \boldsymbol{\varepsilon}_t \tag{6}$$

This SSF gives the setup we have described when $\mathbf{T}_3 = \mathbf{0}$ and $\mathbf{T}_4 = \mathbf{0}$, but otherwise allows for what Jacobs and van Norden (2006) refer to as 'spillover effects'.

3.1 Implications of the SS Model of data revisions for the Rodriquez Mora and Schulstad (2007) regression estimates

Given this statistical model of data revisions, we can calculate the least-squares population values of the coefficients in the regression (2) for both news and noise revisions. We begin by assuming the absence of spillovers ($\mathbf{T}_3 = \mathbf{0}$ and $\mathbf{T}_4 = \mathbf{0}$) and that $v_t^{t+l} = \varepsilon_t^{t+l} = 0$, so that $y_t^{t+l} = y_t$. That is, we assume that the series taken as final values are equal to the true values. Our findings for this case are recorded in the following proposition (with detailed proofs confined to the appendix).

Proposition 1. When the revisions process is given by the measurement and transition equations in (6) and (5), and assuming that the final values are the true values, $v_t^{t+l} = \varepsilon_t^{t+l} = 0$, then the population values of the parameters in the regression model (2) are $\alpha_1 = \rho$ and $\alpha_2 = 0$ under both pure news and pure noise data revisions processes in the absence of spillovers.

Suppose there are news and noise spillovers of the type that $C\left(v_t^{t+s}, v_{t-1}^{t+s-1}\right) = \phi V\left(v_{t-1}^{t+s-1}\right)$ and $C\left(\varepsilon_t^{t+s}, \varepsilon_{t-1}^{t+s-1}\right) = \delta V\left(\varepsilon_{t-1}^{t+s-1}\right)$ for $s = 1, \ldots, l$. This suggests that the extent of the upward revision to one period (say, t) is related to the extent of the upward revision to the previous period $(t-1)$. In other words, the revisions are serial correlated. If we allow for this type of spillover, then the true process for y_t in the SS model (eq. 3) needs to be amended to

$$y_t = \rho_1 y_{t-1} + R_1 \eta_{1t} + \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i} - \phi v_{t-1}^{t},$$

to reflect the fact that

$$v_t^{t+1} = \phi v_{t-1}^{t} - \sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i}.$$

The results for when we allow spillovers of this sort are collected in the following proposition.

Proposition 2. When the revisions process is given by the measurement and transition equations in (6) and (5), while news and noise spillovers are allowed, and maintaining the assumption that $v_t^{t+l} = \varepsilon_t^{t+l} = 0$, the population values of α_1 and α_2 in (2) are $\alpha_1 = \rho + \phi$ and $\alpha_2 = -\phi$ for news revisions, and $\alpha_1 = \rho$ and $\alpha_2 = 0$ for noise revisions.

Thus the coefficient on the lagged first-announcement, α_2, equals zero when there are noise spillovers, but is non-zero in the presence of news spillovers. Moreover, in the case of news spillovers $\alpha_1 + a_2 = \rho$.

Suppose now that the final series y_t^{t+l} does not reveal the truth, as either $v_t^{t+l} \neq 0$ or $\varepsilon_t^{t+l} \neq 0$. We collect our results in two propositions, treating the case of noise and news revisions separately.

Proposition 3. When the revisions process is given by the measurement and transition equations in (6) and (5) but with only noise revisions, and when in addition $\varepsilon_t^{t+l} \neq 0$, in general $\alpha_1 \neq \rho$ and $\alpha_2 \neq 0$, and the values of α_1 and α_2 depend on the properties of the revisions process. However, we can establish a number of interesting special cases, notably, i) if $\rho = \delta$, then $\alpha_1 = \rho$ and $\alpha_2 = 0$; ii) if $\rho > 0$, $\delta > 0$, and $\rho > \delta$, then $0 < \alpha_2 < \rho - \delta$, and iii) if $\delta = 0$, then $\alpha_1/\alpha_2 = V_{\varepsilon_1}/V_{\varepsilon_l}$, the ratio of the variances of the first and last noise components.

Proposition 4. When the revisions process is given by the measurement and transition equations in (6) and (5) but with only news revisions, and when in addition $v_t^{t+l} \neq 0$, then $\alpha_1 = \rho$ and $\alpha_2 = 0$

when $\phi = 0$. When $\phi \neq 0$, $\alpha_1 \approx \rho + \phi$ and $\alpha_2 \approx -\phi$, where the approximations depend on ϕ being small.

Hence, if the SSM is a good model of the data revisions process, we would expect to find α_2 non-zero in (2) when either: revisions are noisy, and the series of final observations remain noisy estimates of the true values, i.e., $\varepsilon_t^{t+l} \neq 0$; or, revisions are news, and the news revisions are correlated (irrespective of whether the final series reveal the true values).

3.2 Empirical Assessment of Analytical Results

We have established that the characterisation of revisions as news or noise may affect whether announcements matter, in the sense that $\alpha_2 \neq 0$ in regression (2). Revisions are defined as reducing noise if the initial estimate is an observation on the final series but measured with error, in which case the revisions are uncorrelated with revised data, but typically are correlated with data available when the initial estimate was made. Hence noisy revisions are predictable. Alternatively, revisions contain news if the initial estimate is an efficient forecast of the revised data or final value, such that the revision is unpredictable from information available at the time the initial estimate was made.[3] We test for news and noise revisions using, respectively, the following regressions:

$$y_t^f - y_t^{t+1} = \alpha + \beta y_t^{t+1} + \omega_{1t} \quad (7)$$

$$y_t^f - y_t^{t+1} = \alpha + \beta y_t^f + \omega_{2t} \quad (8)$$

where the null hypothesis is that $\alpha = \beta = 0$ in both cases. y_t^f is one of the three estimates of the final value, y_t^{t+12}, y_t^{t+24}, y_t^{2008Q2}, and we also include tests of the revisions between y_t^{t+12} and y_t^{t+24}, and between y_t^{t+24} and y_t^{2008Q2}.

Table 2 summarises our results. Focusing firstly on revisions that include the initial data (y_t^{t+1}), a key finding is that the revisions from the initial data to the 3 and 6 year vintages (y_t^{t+12}, y_t^{t+24}) appear to be characterised as news (although the news hypothesis would be rejected at the 10% level for the longer-period revision). The news hypothesis is clearly rejected for the revision to the 2008Q2 vintage. For the revision from the 6-year vintage to the 2008Q2 vintage we also reject the news hypothesis. In nearly all cases, and in all cases of revisions involving first announcements, the noise hypothesis is rejected. In summary, it appears that revisions between the first release and three years later can be characterised as containing news, whereas longer period revisions are more difficult to characterise.

This pattern of results suggests that the nature of the revisions process is influenced by benchmark revisions. Generally, the first three years of revisions are based on new information and enhance the estimates, and consequently show up as news revisions, whereas revisions after three

[3] See Mankiw and Shapiro (1986) for an early contribution: they found that revisions to real output added news. See also Croushore and Stark (2003) and Aruoba (2008).

years are essentially benchmark revisions of a methodological nature which provide little new information on more distant past observations.[4] For example, the base year changes that formed part of the benchmark changes in 1976, 1985 and 1991 disproportionately affected more distant observations under the fixed-weighting scheme and are likely to have contributed to the rejection of the news hypothesis when revisions are calculated using the 2008Q2 vintage data.

We use these findings to guide the specification of our empirical SS model. However, it is important to appreciate at the outset that the SS model is not designed to capture the structural changes that constitute the benchmark revisions, and it is not clear how this could best be done (but van Dijk *et al.* (2007) offer a promising approach). We would hope to be able to model the regular revisions that characterise the data for up to three years, but doubt that such a model will prove adequate for the 2008Q2 vintage.

We make a number of adjustments and simplifications to the general framework of section 3 to facilitate the estimation of the SSM, and to more accurately estimate the parameters that the analytical results in section 3.1 suggest are key to determining the apparent relationship between real growth and early announcements. The population values of the parameters α_1 and α_2 in (2) depend on the process for the true data, y_t, on the nature of the revisions (news, noise; spillovers), and on whether y_t is observable (that is, whether $y_t^{t+l} = y_t$). When y_t is only observable with error ($\sigma_{\varepsilon_l} \neq 0$), the size of the measurement error also affects the values of α_1 and α_2. Based on table 2, it appears we can set $\sigma_{\varepsilon_1} = \sigma_{\varepsilon_2} = \ldots = \sigma_{\varepsilon_l} = 0$. Estimating the variance of the news revisions turned out to be difficult unless we assume that $\sigma_{v_l} = 0$. [5] Setting $\sigma_{v_l} = 0$ implies that $y_t^{t+l} = y_t$ in the absence of noise, so a pragmatic solution is simply to allow measurement error only for the last vintage (y_t^{t+l}).

Although we have data vintages from $t+1$ to $t+24$, and 2008Q2, we select only some vintages to use in the estimation of the model. This keeps parameter proliferation in check while providing enough information to accurately estimate the parameters of interest. Specifically, we use:

$$\mathbf{y}_t = \left(y_t^{t+1}, y_t^{t+4}, y_t^{t+8}, y_t^{t+12}, y_t^{t+24}, y_t^{08:Q2} \right)'.$$

Given the above specification of the revision process, the state vector is:

$$\boldsymbol{\beta}_t = \left(y_t, y_{t-1}, v_t^{t+1}, v_t^{t+4}, v_t^{t+8}, v_t^{t+12}, v_t^{t+24}, v_t^{08:Q2}, \varepsilon_t^{08:Q2} \right)',$$

where we assume $\sigma_{v_{08:Q2}} = 0$. The data suggest that the news revisions are serially correlated - the

[4] Aruoba (2008) takes the view that three years of revisions account for the information content of the revisions to US macro variables.

[5] σ_{v_l} appears to be poorly identified. Jacobs and van Norden (2006, Table 3) appear to find the same: their estimate of σ_{v_l} (also when only allowing news) is ten times larger than the other news variances.

spillover phenomenon described in section 3.1. The $\mathbf{T}$ matrix of equation (5) is specified as:

$$\mathbf{T} = \begin{bmatrix} \rho & 0 & -\phi & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \phi & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \phi & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \phi & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \phi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \phi & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \phi & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We also add a constant to the y_t equation. Figure 1 present smoothed estimates of the state variables including confidence intervals using data from 1975:Q3. The crucial estimates are $\hat{\rho} = .41$, $\hat{\phi} = -.11$ and $\sigma^2_{\varepsilon_t^{08:Q2}} = .048$, all of which are statistically significant at the 5% level.

From the estimated parameters of the SSM we can deduce the values of α_1 and α_2 in (2). As a check on the usefulness of the analytical results for small samples, we simulate data from the estimated SSM and calculate the Monte Carlo estimates (as averages over replications) of quantities of interest, such as the parameter estimates and rejection frequencies of tests of the significance of the parameters. The results reported in table 3 are based on two sample sizes, $T = 150$ and $T = 1,000$. In the first case the researcher has just over 37 years of quarterly data. For each simulated series of size T from the data generating process, the estimated SSM, we estimate equation (2) and record whether the null hypotheses that $\alpha_1 = 0$ and that $\alpha_2 = 0$ are rejected at the 5% level using t-tests with robust standard errors. Table 3 presents the average estimates over replications and the proportion of replications that the null hypothesis was rejected assuming different vintages as final data (all vintages incorporated in the state-space model).

Consider firstly the two regressions based on using y^{t+24} and y^{t+12} respectively as the final values, and y^t_{t-1} as the second explanatory variable, when T takes on the empirically relevant value of 150. Given the specification of the revisions process, i.e., news spillovers, Proposition 4 predicts that $\alpha_1 = \rho + \phi = .41 - .11 = .30$, and $\alpha_2 = -\phi = .11$, which is a close match to the Monte Carlo estimates of these quantities. The Monte Carlo point estimates are 0.293 for α_1 for both y^{t+24} and y^{t+12}, and 0.099 and 0.098 respectively for α_2. The simple analytical formulae do not apply directly when there is measurement error, as when y_t^{2008Q2} are used as final values, but nevertheless the clear finding is that lagged output is more important than the lagged first announcement, both in terms of magnitude, and in terms of its statistical significance. For both y^{t+24} and y^{t+12}, the rejection frequency of $\alpha_2 = 0$ is only around 0.09, while that for $a_1 = 0$ is around one half. If instead we set $T = 1,000$ the rejection frequency for $\alpha_1 = 0$ is approximately one, and that for $\alpha_2 = 0$ increases, but the relative magnitudes of the sizes of the estimated effects of lagged final output and lagged first announcements is clearly at odds with the findings of Rodriquez Mora and Schulstad (2007), and with our empirical findings when y^{t+24} and y^{2008Q2} are used as finals.

The key finding is that the SSM is only able to capture the aspects of the data revisions process that relate to 'regular revisions'. The y_t^{t+12} vintage consists mainly of regular revisions to y_t^{t+1}. Hence the Monte Carlo estimates of the regression (2) based on the estimated SSM match the empirical findings when we consider the relationship between y_t^{t+12}, and y_{t-1}^{t+11} and y_{t-1}^{t}. The SSM does not capture the structural aspects of the benchmark revisions that generate the empirical correlation between y_t^{t+24} and y_{t-1}^{t} (or between y_t^{2008Q2} and y_{t-1}^{t}) so that the simulation results for this case do not match the empirical findings, and are similar to those for y_t^{t+12}. In principle the SSM is capable of generating the finding that the first announcement affects output growth, but this would require a news spillover (ϕ) over twice as large as the empirical estimate.

4 A vector autoregressive model of data vintages

The vector autoregression (VAR) is an alternative to the SSM for modelling different data vintages (see e.g., Garratt, Lee, Mise and Shields (2006)). Whereas the news and noise components in the SSM are unobserved, the VAR models the relationship between observables directly without recourse to additional latent variables. This allows the use of models of data revisions with more complex dynamics. Models which allow for potential instabilities caused by benchmark revisions have been used in related contexts, such as the flexible time-varying parameter (TVP) model of van Dijk *et al.* (2007), but for our current purposes the constant-parameter VAR will be used. We require a systems model of the set of different data estimates in order to be able to simulate, whereas the TVP model of van Dijk *et al.* (2007) is a single-equation forecasting model.

To motivate our approach, suppose that there are just two vintages, and the second vintage (i.e., the first revision) reveals the truth (so $y_{t-2}^{t} = y_{t-2}$). The vector of endogenous variables consists of the new information that becomes available at period t, $Y_t = \left(y_{t-1}^{t}, y_{t-2}^{t}\right)'$. At period t we also observe y_{t-3}^{t}, $y_{t-4}^{t}, \ldots$, etc., but under the assumption that the second vintage reveals the truth, $y_{t-3}^{t} = y_{t-3}^{t-1} = y_{t-3}$, which was in the period $t-1$ data set. As we are modelling growth rates, it is reasonable to assume that the elements of Y_t are stationary so that Y_t has a Wold representation. We assume that this can be approximated by a finite order VAR of order p:

$$Y_t = \Gamma_0 + \sum_{i=1}^{p} \Gamma_i Y_{t-i} + \varepsilon_t$$

where $Y_{t-i} = \left(y_{t-1-i}^{t-i}, y_{t-2-i}^{t-i}\right)$. $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})'$ are the innovations to the variables $\left(y_{t-1}^{t}, y_{t-2}^{t}\right)$ based on the information set comprising $I_{t-1} = \{Y_{t-1}, Y_{t-2}, \ldots\}$, and in general will be correlated, so that $\Omega = E\left(\varepsilon_t \varepsilon_t'\right)$ is non-diagonal.

Within this framework, a more general version of the 'news hypothesis' can be tested and applied as a restriction on the VAR if warranted. The simple news hypothesis is often implemented as a test of $E\left(y_t^{t+2} - y_t^{t+1} \mid y_t^{t+1}\right) = 0$, i.e., the revision to the first estimate is unrelated to the first estimate. The more general version is $E\left(y_t^{t+2} - y_t^{t+1} \mid I_{t+1}\right) = 0$, so that the revision does not vary systematically with past values of Y_{t+1}, $Y_t, \ldots$. In terms of the VAR this requires that $\gamma_2 = 0$, $\gamma_{1;2,1} = 1$, $\gamma_{1;2,2} = 0$ and $\gamma_{i;2,j} = 0$ for all $i > 1$ and $j = 1, 2$, where $\Gamma_0 = (\gamma_1, \gamma_2)'$, and $\gamma_{i;j,k} = \Gamma_{i[j,k]}$ (the j, k element of Γ_i). Hence $y_{t-2}^t = y_{t-2}^{t-1} + \varepsilon_{2t}$ with $E(\varepsilon_{2t} \mid I_{t-1}) = 0$.

As noted in section 2, data revisions are ongoing, but suppose there are q revisions before the truth is revealed. This suggests a $q+1$ dimensional Y_t, $Y_t = \left(y_{t-1}^t, y_{t-2}^t, \ldots y_{t-q-1}^t\right)'$, where $y_{t-q-2}^t = y_{t-q-2}^{t-1}$, so that data on time periods prior to $t-q-1$ are already known in period $t-1$. If we model all q revisions using a p-th order VAR, there will be $p \times (q+1)$ slope coefficients in each of the $q+1$ equations. Our interest is primarily in the variables in regression (2), that is, the first estimate and a measure of 'final' output growth. There would be no loss of information in estimation if we simply estimated the 2-variable system consisting of the VAR equations for y_{t-1}^t and y_{t-q-1}^t. Formally this amounts to estimating the equation system given by:

$$\begin{aligned} S'Y_t &= S'\Gamma_0 + \sum_{i=1}^{p} S'\Gamma_i Y_{t-i} + S\varepsilon_t \\ Y_t^{(2)} &= \Gamma_0^{(2)} + \sum_{i=1}^{p} \Gamma_i^{(2)} Y_{t-i} + \varepsilon_t^{(2)} \end{aligned} \quad (9)$$

where the 2×1 vector $Y_t^{(2)} \equiv S'Y_t$, etc., and S is the $(q+1) \times 2$ selection matrix that picks up the first and last elements of Y_t. However there would be a loss of information in using the estimated model to simulate data as the second through to q-estimates are no longer modelled. We make a further simplification and instead estimate bivariate VARs for $Y_t = \left(y_{t-1}^t, y_{t-q-1}^t\right)'$. Relative to (9), the explanatory variables are lags of $\left\{y_{t-1}^t, y_{t-q-1}^t\right\}$, as opposed to lags of $\left\{y_{t-1}^t, y_{t-2}^t, \ldots, y_{t-q-1}^t\right\}$. The advantages are that we are able to entertain a longer number of lags and the estimated model can be used to simulate data.

4.1 Specification issues and empirical results

We estimate two VAR models to investigate the dynamic relationships between the first and final data. In the first, $\mathrm{VAR}_{q=12}$, the final data were taken to be the data estimates made three years ($q = 12$) after the time period to which the observation refers, and in the second case, $\mathrm{VAR}_{q=24}$, the final data are the estimates after six years ($q = 24$). For the first VAR we set $p = 12$, and for the second $p = 24$.

Consider $\mathrm{VAR}_{q=12}$ with $Y_t = \left(y_{t-1}^t, y_{t-12}^t\right)'$. The second equation in the VAR allows y_{t-12}^t to depend on $\left\{y_{t-2}^{t-1}, \ldots, y_{t-13}^{t-12}; y_{t-13}^{t-1}, \ldots, y_{t-24}^{t-12}\right\}$. That is, the period t estimate of $t-12$ (assumed to be the final value of y in $t-12$) can depend on estimates of more recent observations (say, y at time $t-2$) provided those estimates are vintage-dated earlier than t. The choice of $q = 12$ reflects the fact that regular revisions typically occur for the first three years, as discussed in section 2. For

VAR$_{q=12}$ the news hypothesis (i.e., $E\left(y_{t-12}^{t} - y_{t-12}^{t-11} \mid I_{t-11}\right) = 0$) requires that the coefficient on y_{t-12}^{t-11} in the equation for y_{t-12}^{t} is unity, and the coefficients on $\left\{y_{t-13}^{t-12},\, y_{t-23}^{t-11},\, y_{t-24}^{t-12}\right\}$ in the equation for y_{t-12}^{t} are all zero.

VAR$_{q=24}$ has $Y_t = \left(y_{t-1}^{t}, y_{t-24}^{t}\right)'$ and $p = 24$. This is to approximate taking the final values from the last available data vintage. Hence these final data will contain benchmark revisions as well as regular revisions. Using $\left\{y_t^{208Q2}\right\}$ as final values would be inconsistent with the interpretation of the elements of the vector Y_t as the new information that is revealed at period t, as the latest vintage values are not revealed until $t = 2008Q2$.[6] We also need to increase the lag length for the second VAR if we wish to allow the news hypothesis to be incorporated as a special case. The news hypothesis requires that $E\left(y_{t-24}^{t} - y_{t-24}^{t-23} \mid I_{t-23}\right) = 0$, so that the VAR needs to include y_{t-24}^{t-23}, the 23rd lag of y_{t-1}^{t} (the first element of Y_t).

We estimated the two VARs on the common sample period 1977:3 to 2004:4. We then simulated data from the estimated VARs, and ran the Rodriquez Mora and Schulstad (2007) regression, recording the Monte Carlo estimates of the estimated parameters as well as the rejection frequencies of the standard t-tests of the hypotheses of the individual insignificance of the two explanatory variables. The model disturbances were drawn from a bivariate normal distribution with covariance structure equal to that of the VAR residuals. We simulate $T + 300$ observations, and discard the initial 300 observations to minimise the impact of initial values. As in the case of the SSM, of interest is whether the VAR is able to reproduce the underlying data generation process sufficiently well that the empirical findings hold in the simulated data.

We also estimated both VARs models imposing the restrictions of news revisions. For both cases $q = 12$, 24, we are not able to reject the restrictions imposed by the news hypothesis using a Wald statistic (details available on request). However, imposing the news hypothesis does have a significant impact on some of our simulation results, as we explain below. This apparent anomaly - the restricted and unrestricted models used as data generating processes yield marked differences in terms of estimates of equation (2), even though one is a valid restriction of the other - we attribute to the relatively large number of model parameters compared to the available number of observations, resulting in the imposition of the restriction having only a small effect on the maximised value of the model likelihood function.

We firstly consider the results when the news restriction is not imposed. The results are recorded in the first two columns of table 4 for $T = 150, 1,000$. When $T = 150$, we find that the magnitude of the effect of the first announcement is larger than that of lagged growth in the first VAR (VAR$_{q=24}$, using y^{t+24} as finals), and the first announcement is more often significant. In the second VAR (VAR$_{q=12}$) for y^{t+12} the situation is reversed. The lagged final value has the larger coefficient and is more often significant. For the larger sample size the rejection frequencies accentuate the

[6] Formally, we could write $Y_t = \left(y_{t-1}^{t}, y_{t-\infty}^{t}\right)'$ for all t prior to 2008$Q2$, as for these t no final values are revealed.

importance of the first announcement for y^{t+24}, and the importance of the lagged final value for y^{t+12}. These findings are in tune with the empirical estimates recorded in table 1[7]: especially that the significance of the first announcement in (2) depends on whether the final series contains benchmark revisions (y^{t+24}) or not (y^{t+12}). The use of long lags in the VAR (especially $\text{VAR}_{q=24}$) results in a model which is better able to capture the effects of benchmark revisions than the SSM: the $\text{VAR}_{q=24}$ adequately captures the relationship between the initial data and that revised six years later.

Consider now the impact of imposing the news hypothesis on the VAR models, and using the restricted VAR models in the Monte Carlo. The results are recorded in the final two columns of table 4. The results for $\text{VAR}_{q=24}$ are little changed. Recall from table 2 that we failed to reject both the news and noise hypotheses for y^{t+24}. For y^{t+12} the effect of imposing the news hypothesis is to push the estimated value of α_1 close to zero and double the size of α_2, pointing to the importance of first announcements in (2). This is markedly at odds with the empirical findings for y^{t+12} reported in table 1, and suggests that the characterisation of revisions between $t+1$ and $t+12$ as being purely news is too simplistic: in table 2 we found that the news hypothesis was borderline significant at the 10% level. This suggests a more complex dynamic process for revisions than that they are simply news, and that the unrestricted VAR captures this reasonably well.

We also use a VAR with $q = 12$ to model the revision processes of y^t_{t-12} and y^t_{t-24}, and the simulation results with this data-generating process are presented in the last two columns of table 4. The imposition of the news hypothesis has a minor effect on the estimates, in agreement with the results of table 2. Comparing the estimates of equation (2) using the VAR data generating process and the results using the actual data in the last column of table 1, we conclude that in general the VAR model with the imposition of the news hypothesis is a good representation of the revision process from the $t + 12$ up to the $t + 24$ vintage. Between these two vintages, the lagged value of the $t + 12$ vintage affects y^{t+24} but the lagged 'final' value (y_{t-1}) is insignificant.

5 Conclusions

In the recent literature the correlation between real output growth and the first announcement of growth in the previous period (controlling for lagged real output growth) has been given a

[7]For example, in table 4 we obtain Monte Carlo estimates of α_2 of 0.221 and 0.126 respectively for y^{t+24} and y^{t+12}, compared to empirical estimates in table 1 of 0.238 and 0.124 for the comparable sample period (1974–2001). The Monte Carlo results also indicate that whereas $a_2 = 0$ is only rejected a third of the time for y^{t+24} when $T = 150$, the rejection frequency is close to one when $T = 1,000$, suggesting that the lack of significance we find empirically might be due to the small sample size.

behavioural interpretation: agents respond to the first announcement, and this affects the future course of real output growth. We show that instead this correlation can be viewed as a by-product of the data revision process, and in particular, is occasioned by the structural changes that accompany benchmark revisions. The importance of the first announcement is found to depend crucially on whether the series taken as the final estimates of real output contains primarily regular revisions that add news to the initial estimate, or also includes benchmark revisions.

We analyse whether two popular modelling frameworks for data revisions are able to adequately model the relationship between the initial and 'final' data to the extent that data simulated from these models matches the empirical data in certain respects. Namely, first announcements do not have predictive power for three-year revised final values, but do appear to explain six-year revised finals. The state-space modelling framework that characterises data revisions as (unobservable) news and noise processes affecting a (generally unobserved) true value does not match the empirical findings, although our analytical results suggest that it would do so were revisions news with a larger spillover effect. That is, the first-announcement of output growth would help predict future output growth even controlling for lagged output growth if data revisions add new information to the measurement of output growth and are negatively serial correlated. In contrast, a vector-autoregression in terms of observables provides a closer match to the empirical findings for the six-year revised final series, but is unable to explain the empirical findings for the three-year revised final series if we assume that revisions are purely news within this framework.

Given the high degree of collinearity between the two explanatory variables, the first announcement of the lagged value, and the 'final' lagged value, of output growth, and the complexity of the revision process between the first and final value, it is perhaps unsurprising that the relative importance of the two explanatory variables depends on the definition of 'final'. The two modelling approaches we consider help to partially illuminate some aspects of the problem, but neither serves as a comprehensive statistical framework. However taken together they are sufficiently informative to indicate that the presumption that the first announcement has a causal impact on future output growth appears unwarranted.

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. (2008a). Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *Journal of Applied Econometrics*. Forthcoming.

Clements, M. P., and Galvão, A. B. (2008b). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States. *Journal of Business and Economic Statistics*, **26**, 546–554. No. 4.

Corradi, V., Fernandez, A., and Swanson, N. R. (2007). Information in the revision process of real-time datasets. Working paper, Rutgers University.

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.

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

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.

Garratt, A., Lee, K., Mise, E., and Shields, K. (2006). Real time representations of the output gap. mimeo, University of Leicester.

Hall, R. E. (1978). Stochastic implications of the life cycle-permanent income hypothesis: Evidence. *Journal of Political Economy*, **86**, 971–987.

Jacobs, J. P. A. M., and van Norden, S. (2006). Modeling data revisions: Measurement error and dynamics of 'true' values. Ccso working paper 2006/07, CCSO Centre for Economic Research, University of Groningen.

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.

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.

Oh, S., and Waldman, M. (1990). The macroeconomic effects of false announcements. *The Quarterly*

Journal of Economics, **105**, 1017–1034. No. 4.

Oh, S., and Waldman, M. (2005). The index of leading economic indicators as a source of expectational shocks. *Eastern Economic Journal*, **31**, 75–95. No. 1.

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.

Rodriquez Mora, J. V., and Schulstad, P. (2007). The effect of GNP announcements on fluctuations of GNP growth. *European Economic Review*, **51**, 1922–1940.

Runkle, D. E. (1998). Revisionist history: How data revisions distort economic policy research. *Federal Reserve Bank of Minneapolis Quarterly Review*, **22**, 3–12. No. 4.

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.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, **48**, 1–48.

van Dijk, D., Franses, P. H., and Ravazzolo, F. (2007). Evaluating real-time forecasts in real-time. Mimeo, Econometric Institute Research Report EI2007-33, Erasmas University, Rotterdam.

A Proofs of Propositions

For the state space data generating process outlined in section 3, we derive the population values of α_1 and α_2 in equation (2).

In the general case we allow that $y_t^{t+l} \neq y_t$, because either $\sigma_{v_l} \neq 0$ and/or $\sigma_{\varepsilon_l} \neq 0$.

From section 3 we have that:

$y_t^{t+1} = y_t + v_t^{t+1} + \varepsilon_t^{t+1}$, $y_t^{t+l} = y_t + v_t^{t+l} + \varepsilon_t^{t+l}$ and $y_t = \rho y_{t-1} + R_1 \eta_{1t} - v_t^{t+1}$.

Allowing for news spillovers, $v_t^{t+1} = \phi v_{t-1}^t - \sum_{i=1}^l \sigma_{v_i} \eta_{2t,i}$, $v_t^{t+l} = \phi v_{t-1}^{t+l-1} - \sigma_{v_l} \eta_{2t,l}$; and for noise spillovers, $Cov\left(\varepsilon_t^{t+1}, \varepsilon_{t-1}^t\right) = \delta V\left(\varepsilon_{t-1}^t\right)$.

The population values of α_1 and α_2 in the equation (2) are given by:

$$\alpha_1 = \frac{V(y_{t-1}^t)C(y_t^{t+l}, y_{t-1}^{t+l-1}) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)C(y_t^{t+l}, y_{t-1}^t)}{V(y_{t-1}^{t+l-1})V(y_{t-1}^t) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)^2}.$$

$$\alpha_2 = \frac{V(y_{t-1}^{t+l-1})C(y_t^{t+l}, y_{t-1}^t) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)C(y_t^{t+l}, y_{t-1}^{t+l-1})}{V(y_{t-1}^{t+l-1})V(y_{t-1}^t) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)^2}.$$

To compute these values one needs to compute the required moments: (1) $C(y_t^{t+l}, y_{t-1}^t)$; (2) $C(y_{t-1}^{t+l-1}, y_{t-1}^t)$; (3) $C(y_t^{t+l}, y_{t-1}^{t+l-1})$; (4) $V(y_{t-1}^t)$, and (5) $V(y_{t-1}^{t+l-1})$. To compute these moments, we use the following notation:

$$\begin{aligned}
V_y &= V\left(y_t\right); \\
V_{v_1} &= V\left(\sum_{i=1}^{l} \sigma_{v_i} \eta_{2t,i}\right), V_{v_l} = V\left(\sigma_{v_l} \eta_{2t,l}\right); \\
V_{1,l} &= C\left(v_t^{t+1} v_t^{t+l}\right), V_{11} = C\left(v_t^{t+1} v_t^{t+1}\right); \\
V_{\varepsilon_1} &= V\left(\varepsilon_{t-1}^t\right), V_{\varepsilon_l} = V\left(\varepsilon_{t-1}^{t+l-1}\right).
\end{aligned}$$

(1A)

$$\begin{aligned}
C(y_t^{t+l}, y_{t-1}^t) &= Cov\left[\left(y_t + v_t^{t+l} + \varepsilon_t^{t+l}\right)\left(y_{t-1} + v_{t-1}^t + \varepsilon_{t-1}^t\right)\right] \\
&= Cov\left(y_t y_{t-1}\right) + Cov\left(y_t v_{t-1}^t\right) + Cov\left(y_{t-1} v_t^{t+l}\right) + C\left(v_t^{t+l} v_{t-1}^t\right) \\
&= \rho V_y + (\rho - \phi)\, Cov\left(y_{t-1} v_{t-1}^t\right) - \phi V_{11} + \phi Cov\left(y_{t-1}, v_{t-1}^{t+l-1}\right) + \phi V_{1,l}
\end{aligned}$$

(2A)

$$\begin{aligned}
C(y_{t-1}^{t+l-1}, y_{t-1}^t) &= Cov\left[\left(y_{t-1} + v_{t-1}^{t+l-1} + \varepsilon_{t-1}^{t+l-1}\right)\left(y_{t-1} + v_{t-1}^t + \varepsilon_{t-1}^t\right)\right] \\
&= V\left(y_{t-1}\right) + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + Cov\left(v_{t-1}^{t+l-1}, v_{t-1}^t\right) \\
&= V_y + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + V_{1,l}
\end{aligned}$$

(3A)

$$
\begin{aligned}
C(y_t^{t+l}, y_{t-1}^{t+l-1}) &= C\left[\left(y_t + v_t^{t+l} + \varepsilon_t^{t+l}\right)\left(y_{t-1} + v_{t-1}^{t+l-1} + \varepsilon_{t-1}^{t+l-1}\right)\right] \\
&= Cov\left(y_t y_{t-1}\right) + C\left(y_t v_{t-1}^{t+l-1}\right) + C\left(y_{t-1} v_t^{t+l}\right) + C\left(v_t^{t+l} v_{t-1}^{t+l-1}\right) + C\left(\varepsilon_t^{t+l} \varepsilon_{t-1}^{t+l-1}\right) \\
&= \rho V\left(y_{t-1}\right) - \phi C\left(y_{t-1} v_{t-1}^{t}\right) + \rho C\left(y_{t-1} v_{t-1}^{t+l-1}\right) - C\left(v_t^{t+1} v_{t-1}^{t+l-1}\right) + \\
&\quad + \phi C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \phi V\left(v_{t-1}^{t+l-1}\right) + \delta V\left(\varepsilon_{t-1}^{t+l-1}\right) \\
&= \rho V\left(y_{t-1}\right) - \phi C\left(y_{t-1} v_{t-1}^{t}\right) + (\rho + \phi) C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \delta V\left(\varepsilon_{t-1}^{t+l-1}\right) \\
&= \rho V_y + (\rho + \phi) C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \delta V_{\varepsilon_l} - \phi C\left(y_{t-1}, v_{t-1}^{t}\right)
\end{aligned}
$$

(4A)

$$
\begin{aligned}
V\left(y_{t-1}^{t}\right) &= V\left(y_{t-1}\right) + V\left(v_{t-1}^{t}\right) + V\left(\varepsilon_{t-1}^{t}\right) + 2C\left(y_{t-1} v_{t-1}^{t}\right) \\
&= V_y + V_{11} + V_{\varepsilon_1} + 2C\left(y_{t-1} v_{t-1}^{t}\right)
\end{aligned}
$$

(5A)

$$
\begin{aligned}
V\left(y_{t-1}^{t+l-1}\right) &= V\left(y_{t-1}\right) + V\left(v_{t-1}^{t+l-1}\right) + V\left(\varepsilon_{t-1}^{t+l-1}\right) + 2C\left(y_{t-1} v_{t-1}^{t+l-1}\right) \\
&= V_y + V_{1,l} + V_{\varepsilon_l} + 2C\left(y_{t-1} v_{t-1}^{t+l-1}\right)
\end{aligned}
$$

All these moments can be further simplified by using that:

$$
V_{1,l} = V_{v_l}\left(1 - \phi^2\right)^{-1}
$$
$$
V_{11} = V_{v_1}\left(1 - \phi^2\right)^{-1}.
$$

$$
C\left(y_{t-1} v_{t-1}^{t+l-1}\right) = \rho\phi C\left(y_{t-2} v_{t-2}^{t+l-2}\right) - V_{1l} = -V_{1l}(1-\rho\phi)^{-1} = -V_{v_l}(1-\rho\phi)^{-1}(1-\phi^2)^{-1}.
$$

$$
C\left(y_{t-1} v_{t-1}^{t}\right) = -V_{11}(1-\rho\phi)^{-1} = -V_{v_1}(1-\rho\phi)^{-1}(1-\phi^2)^{-1}.
$$

Re-writing using previous results:

(1B)

$$
\begin{aligned}
C(y_t^{t+l}, y_{t-1}^{t}) &= \rho V_y + (\rho - \phi) Cov\left(y_{t-1} v_{t-1}^{t}\right) - \phi V_{11} + \phi Cov\left(y_{t-1}, v_{t-1}^{t+l-1}\right) + \phi V_{1,l} \\
&= \left[V_y \rho (1 - \rho\phi) - \rho\left(1 - \phi^2\right) V_{11} - \rho\phi^2 V_{1,l}\right](1 - \rho\phi)^{-1}
\end{aligned}
$$

(2B)

$$
\begin{aligned}
C(y_{t-1}^{t+l-1}, y_{t-1}^{t}) &= V_y + C\left(y_{t-1} v_{t-1}^{t}\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + V_{1,l} \\
&= \left[V_y (1 - \rho\phi) - V_{11} - \rho\phi V_{1,l}\right](1 - \rho\phi)^{-1}
\end{aligned}
$$

(3B)

$$\begin{aligned} C(y_t^{t+l}, y_{t-1}^{t+l-1}) &= \rho V_y + (\rho+\phi) C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \delta V_{\varepsilon_l} - \phi C\left(y_{t-1}, v_{t-1}^t\right) \\ &= \left[V_y \rho (1-\rho\phi) + \delta V_{\varepsilon_l}(1-\rho\phi) + \phi V_{11} - (\rho+\phi) V_{1,l}\right](1-\rho\phi)^{-1} \end{aligned}$$

(4B)

$$\begin{aligned} V\left(y_{t-1}^t\right) &= V_y + V_{11} + V_{\varepsilon_1} + 2C\left(y_{t-1} v_{t-1}^t\right) \\ &= \left[V_y(1-\rho\phi) + V_{\varepsilon_1}(1-\rho\phi) - V_{11}(1-\rho\phi)\right](1-\rho\phi)^{-1} \end{aligned}$$

(5B)

$$\begin{aligned} V\left(y_{t-1}^{t+l-1}\right) &= V_y + V_{1,l} + V_{\varepsilon_l} + 2C\left(y_{t-1} v_{t-1}^{t+l-1}\right) \\ &= \left[V_y(1-\rho\phi) + V_{\varepsilon_l}(1-\rho\phi) - V_{1,l}(1-\rho\phi)\right](1-\rho\phi)^{-1} \end{aligned}$$

Using 5A, 3A, 2A and 1A, the numerator for α_1 is:

$$\begin{aligned} & V(y_{t-1}^t) C(y_t^{t+l}, y_{t-1}^{t+l-1}) - C(y_{t-1}^{t+l-1}, y_{t-1}^t) C(y_t^{t+l}, y_{t-1}^t) \qquad (10) \\ = & \left[V_y + V_{11} + V_{\varepsilon_1} + 2C\left(y_{t-1} v_{t-1}^t\right)\right] \\ & \times \left[\rho V_y + (\rho+\phi) C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \delta V_{\varepsilon_l} - \phi C\left(y_{t-1}, v_{t-1}^t\right)\right] \\ & - \left[V_y + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + V_{1,l}\right] \\ & \times \left[\rho V_y + (\rho-\phi) Cov\left(y_{t-1} v_{t-1}^t\right) - \phi V_{11} + \phi Cov\left(y_{t-1}, v_{t-1}^{t+l-1}\right) + \phi V_{1,l}\right] \end{aligned}$$

Using 5A, 1A, 2A and 3A, the numerator of α_2 is:

$$\begin{aligned} & V(y_{t-1}^{t+l-1}) C(y_t^{t+l}, y_{t-1}^t) - C(y_{t-1}^{t+l-1}, y_{t-1}^t) C(y_t^{t+l}, y_{t-1}^{t+l-1}) \qquad (11) \\ = & \left[V_y + V_{1,l} + V_{\varepsilon_l} + 2C\left(y_{t-1} v_{t-1}^{t+l-1}\right)\right] \\ & \times \left[\rho V_y + (\rho-\phi) Cov\left(y_{t-1} v_{t-1}^t\right) - \phi V_{11} + \phi C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \phi V_{1,l}\right] \\ & - \left[V_y + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + V_{1,l}\right] \\ & \times \left[\rho V_y + (\rho+\phi) C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + \delta V_{\varepsilon_l} - \phi C\left(y_{t-1}, v_{t-1}^t\right)\right]. \end{aligned}$$

And the denominator of both expressions using 5A, 4A and 2A is:

$$\begin{aligned} & V(y_{t-1}^{t+l-1}) V(y_{t-1}^t) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)^2 \qquad (12) \\ = & V(y_{t-1}^{t+l-1}) V(y_{t-1}^t) - \left(V_y + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right)\right)^2 \\ = & V(y_{t-1}^{t+l-1}) V(y_{t-1}^t) - \left[V_y + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right)\right]^2 \\ = & \left[V_y + V_{11} + V_{\varepsilon_1} + 2C\left(y_{t-1} v_{t-1}^t\right)\right] \times \left[V_y + V_{1,l} + V_{\varepsilon_l} + 2C\left(y_{t-1} v_{t-1}^{t+l-1}\right)\right] \\ & - \left[V_y + C\left(y_{t-1} v_{t-1}^t\right) + C\left(y_{t-1} v_{t-1}^{t+l-1}\right) + V_{1,l}\right]^2 \end{aligned}$$

A.1 Proof of proposition 1:

When $\varepsilon_t^{t+l} = v_t^{t+l} = 0$, the numerator of α_1 (from eq.(10)) is:

$$\begin{aligned}
& \left[V_y + V_{11} + V_{\varepsilon_1} + 2C\left(y_{t-1}v_{t-1}^t\right)\right] \times \left[\rho V_y - \phi C\left(y_{t-1}, v_{t-1}^t\right)\right] \\
& - \left[V_y + C\left(y_{t-1}v_{t-1}^t\right)\right] \times \left[\rho V_y + (\rho - \phi)\, Cov\left(y_{t-1}v_{t-1}^t\right) - \phi V_{11}\right] \\
= \;& \rho V_y V_y + \rho V_y V_{11} + \rho V_y V_{\varepsilon_1} + 2\rho V_y C\left(y_{t-1}v_{t-1}^t\right) \\
& -\phi C\left(y_{t-1}, v_{t-1}^t\right) V_y - \phi C\left(y_{t-1}, v_{t-1}^t\right) V_{11} - \phi C\left(y_{t-1}, v_{t-1}^t\right) V_{\varepsilon_1} - \phi C\left(y_{t-1}, v_{t-1}^t\right) 2C\left(y_{t-1}v_{t-1}^t\right) \\
& -\rho V_y^2 - V_y(\rho - \phi)\, Cov\left(y_{t-1}v_{t-1}^t\right) - C\left(y_{t-1}v_{t-1}^t\right)\rho V_y - C\left(y_{t-1}v_{t-1}^t\right)(\rho - \phi)\, Cov\left(y_{t-1}v_{t-1}^t\right) \\
& +\phi V_{11} V_y + \phi V_{11} C\left(y_{t-1}v_{t-1}^t\right) \\
& = (\rho + \phi)\left(V_y V_{11} - C\left(y_{t-1}v_{t-1}^t\right)^2\right) + \rho\left[V_y - \phi C\left(y_{t-1}, v_{t-1}^t\right)\right] V_{\varepsilon_1}
\end{aligned}$$

The denominator (from eq. (12)) is:

$$V_{11}V_y + V_{\varepsilon_1}V_y - C\left(y_{t-1}v_{t-1}^t\right)^2$$

As a consequence, the population value of α_1 assuming $\varepsilon_t^{t+l} = v_t^{t+l} = 0$ is:

$$\alpha_1 = \frac{(\rho + \phi)\left(V_y V_{11} - C\left(y_{t-1}v_{t-1}^t\right)^2\right) + \rho\left[V_y - \phi C\left(y_{t-1}, v_{t-1}^t\right)\right] V_{\varepsilon_1}}{V_{11}V_y + V_{\varepsilon_1}V_y - C\left(y_{t-1}v_{t-1}^t\right)^2} \tag{13}$$

When revisions are pure news, $\alpha_1 = (\rho + \phi)$. When they are either pure noise or news with $\phi = 0$, one has $\alpha_1 = \rho$. This means that when there is no spillovers, that is, $\phi = 0$, for both the cases of news and noise revisions, one has that $\alpha_1 = \rho$.

Under the assumption that $\varepsilon_t^{t+l} = v_t^{t+l} = 0$, the numerator of α_2 (from equation (11)) is:

$$\begin{aligned}
& V(y_{t-1}^{t+l-1})C(y_t^{t+l}, y_{t-1}^t) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)C(y_t^{t+l}, y_{t-1}^{t+l-1}) \\
= \;& V_y \times \left[\rho V_y - \phi V_{11} + (\rho - \phi)\, Cov\left(y_{t-1}v_{t-1}^t\right)\right] \\
& - \left[V_y + C\left(y_{t-1}v_{t-1}^t\right)\right] \times \left[\rho V_y - \phi C\left(y_{t-1}, v_{t-1}^t\right)\right] \\
= \;& -\phi V_{11}V_y + \phi C\left(y_{t-1}, v_{t-1}^t\right)^2.
\end{aligned}$$

And the population value of α_2 assuming $\varepsilon_t^{t+l} = v_t^{t+l} = 0$,

$$\alpha_2 = \frac{-\phi V_{11}V_y + \phi C\left(y_{t-1}v_{t-1}^t\right)^2}{V_{11}V_y + V_{\varepsilon_1}V_y - C\left(y_{t-1}v_{t-1}^t\right)^2}. \tag{14}$$

This implies that the revisions are news $\alpha_2 = \phi$, but if there is no spillovers $\alpha_2 = 0$. When revisions are noise, $\alpha_2 = 0$.

A.2 Proof of Proposition 2:

Using equations (13) and (14), one can show that $\alpha_1 = \rho$ and $\alpha_2 = 0$ when revisions are noise, while $\alpha_1 = \rho + \phi$ and $\alpha_1 = -\phi$ when revisions are news.

A.3 Proof of Proposition 3:

When $\varepsilon_t^{t+l} \neq 0$, $v_t^{t+l} \neq 0$, and there is no spillovers, that is, $\phi = \delta = 0$, the numerator (eq. 10) is:

$$\begin{aligned}
& V(y_{t-1}^t)C(y_t^{t+l}, y_{t-1}^{t+l-1}) - C(y_{t-1}^{t+l-1}, y_{t-1}^t)C(y_t^{t+l}, y_{t-1}^t) \\
= & \left[V_y + V_{11} + V_{\varepsilon_1} + 2C\left(y_{t-1}v_{t-1}^t\right)\right] \times \left[\rho V_y + \rho C\left(y_{t-1}v_{t-1}^{t+l-1}\right)\right] \\
& - \left[V_y + C\left(y_{t-1}v_{t-1}^t\right) + C\left(y_{t-1}v_{t-1}^{t+l-1}\right) + V_{1,l}\right] \times \left[\rho V_y + \rho Cov\left(y_{t-1}v_{t-1}^t\right)\right] \\
= & \rho V_y V_{11} + \rho V_y V_{\varepsilon_1} \\
& + \rho C\left(y_{t-1}v_{t-1}^{t+l-1}\right) V_{11} + \rho C\left(y_{t-1}v_{t-1}^{t+l-1}\right) V_{\varepsilon_1} + \rho C\left(y_{t-1}v_{t-1}^{t+l-1}\right) C\left(y_{t-1}v_{t-1}^t\right) \\
& - \rho V_y V_{1,l} - \rho Cov\left(y_{t-1}v_{t-1}^t\right)^2 - \rho Cov\left(y_{t-1}v_{t-1}^t\right) V_{1,l}
\end{aligned}$$

When $\phi = 0$, one can use the following $V_{1,l} = V_{v_l}$, $V_{11} = V_{v_1}$, $C\left(y_{t-1}v_{t-1}^t\right) = -V_{v_1}$, $C\left(y_{t-1}v_{t-1}^{t+l-1}\right) = -V_{v_l}$. Substituting these expressions gives:

$$\begin{aligned}
= & \rho V_y V_{v_1} + \rho V_y V_{\varepsilon_1} \\
& - \rho V_{v_l} V_{v_1} - \rho V_{v_l} V_{\varepsilon_1} + \rho V_{v_l} V_{v_1} \\
& - \rho V_y V_{v_l} + \rho V_{v_1}^2 + \rho V_{v_1} V_{v_l} \\
= & \rho V_y V_{v_1} + \rho V_y V_{\varepsilon_1} - \rho V_{v_l} V_{\varepsilon_1} - \rho V_y V_{v_l} + \rho V_{v_1}^2 + \rho V_{v_1} V_{v_l}
\end{aligned} \tag{15}$$

Therefore, the population value of numerator of α_1 assuming pure noise and $\varepsilon_t^{t+l} \neq 0$ is:

$$\begin{aligned}
& (V_y + V_{\varepsilon_1}) \times (\rho V_y + \delta V_{\varepsilon_l}) - \rho V_y^2 \\
= & \delta V_y V_{\varepsilon_l} + \rho V_{\varepsilon_1} V_y + \delta V_{\varepsilon_1} V_{\varepsilon_l}
\end{aligned}$$

The denominator (from equation (12)) assuming noise is:

$$\begin{aligned}
& (V_y + V_{\varepsilon_1})(V_y + V_{\varepsilon_l}) - V_y^2 \\
= & V_y V_{\varepsilon_l} + V_{\varepsilon_1} V_y + V_{\varepsilon_1} V_{\varepsilon_l}.
\end{aligned}$$

As a consequence, the population value of α_1 is:

$$\alpha_1 = \frac{\delta V_y V_{\varepsilon_l} + \rho V_{\varepsilon_1} V_y + \delta V_{\varepsilon_1} V_{\varepsilon_l}}{V_y V_{\varepsilon_l} + V_{\varepsilon_1} V_y + V_{\varepsilon_1} V_{\varepsilon_l}}.$$

This means that $\alpha_1 \neq 0$, and depends on the properties of the revision process. And if $\rho = \delta$, then $\alpha_1 = \rho$.

When data revisions are noise, the numerator of α_2 (from equation (11) is:

$$= \rho V_y V_{\varepsilon_l} - V_y \delta V_{\varepsilon_l}$$

The population value of α_2 assuming pure news and $\varepsilon_t^{t+l} \neq 0$ is:

$$\alpha_2 = \frac{(\rho - \delta)\, V_y V_{\varepsilon_l}}{V_y V_{\varepsilon_l} + V_{\varepsilon_1} V_y + V_{\varepsilon_1} V_{\varepsilon_l}}$$

This means that $\alpha_2 \neq 0$. However, if $\rho = \delta$, $\alpha_2 = 0$. In addition if $\rho > 0$, $\delta > 0$, and $\rho > \delta$, then $0 < \alpha_2 < \rho - \delta$. Finally comparing the expressions for α_1 and α_2 in the case of no spillover $(\delta = 0)$, one has that $\alpha_1/\alpha_2 = V_{\varepsilon_1}/V_{\varepsilon_l}$, the ratio of the variances of the first and last noise components.

A.4 Proof of Proposition 4:

When data revisions are pure news and there are no spillovers, the numerator of α_1 (equation (15) becomes:

$$= \rho V_y V_{v_1} - \rho V_y V_{v_l} + \rho V_{v_1}^2 + \rho V_{v_1} V_{v_l}$$

While the denominator (equation (12)) is:

$$\begin{aligned} &= (V_y - V_{v_1})\,(V_y - V_{v_l}) - (V_y - V_{v_1})^2 \\ &= V_y V_{v_1} - V_y V_{v_l} + V_{v_1}^2 + V_{v_1} V_{v_l} \end{aligned}$$

This expression makes use of the fact that $V_{1,l} = V_{v_l}$, $V_{11} = V_{v_1}$, $C\left(y_{t-1} v_{t-1}^{t}\right) = -V_{v_1}$, $C\left(y_{t-1} v_{t-1}^{t+l-1}\right) = -V_{v_l}$ when there are no spillovers. Combining the numerator and the denominator, we have that $\alpha_1 = \rho$.

When data revisions are pure news and there are no spillovers, the numerator of α_2 (equation (11) becomes:

$$\begin{aligned} &= (V_y + V_{v_l} - 2V_{v_l}) \times (\rho V_y - \rho V_{v_1}) \\ &\quad - (V_y - V_{v_1}) \times (\rho V_y - \rho V_{v_l}) \\ &= \rho V_y V_{v_l} - \rho V_y V_{1,l} - \rho V_{v_1} V_{v_l} + \rho V_{v_1} V_{1,l} = 0 \end{aligned}$$

using $C\left(y_{t-1} v_{t-1}^{t+l-1}\right) = -V_{1,l} = -V_{v_l}$ and $C\left(y_{t-1} v_{t-1}^{t}\right) = -V_{11} = -V_{v_1}$.Therefore $\alpha_2 = 0$ when data revisions are pure news and $v_t^{t+l} \neq 0$ with $\phi = 0$.

Now let us assume that there are spillovers, that is, $\phi \neq 0$, when $v_t^{t+l} \neq 0$. The numerator of

α_1 (eq. 10) using expressions 4B, 3B, 2B and 1B and eliminating noise terms is:

$$
\begin{aligned}
&= ([V_y(1-\rho\phi) - V_{11}(1-\rho\phi)] * [V_y\rho(1-\rho\phi) + \phi V_{11} - (\rho+\phi)V_{1,l}]) \\
&\quad -([V_y(1-\rho\phi) - V_{11} - \rho\phi V_{1,l}] * \left[V_y\rho(1-\rho\phi) - \rho\left(1-\phi^2\right)V_{11} - \rho\phi^2 V_{1,l}\right]) \\
&= V_{11}V_y(1-\rho\phi)\phi - V_{11}V_y\rho(1-\rho\phi)(1-\rho\phi) + V_{11}V_y(1-\rho\phi)\left(1-\phi^2\right)\rho + V_{11}V_y\rho(1-\rho\phi) \\
&\quad -V_{1,l}V_y(1-\rho\phi)(\rho+\phi) + V_{1,l}V_y(1-\rho\phi)\rho^2\phi + V_{1,l}V_y\rho^2\phi(1-\rho\phi) \\
&\quad +V_{1,l}V_{11}(1-\rho\phi)(\rho+\phi) - V_{11}V_{1,l}\rho^2\phi^3 - V_{1,l}V_{11}\left(1-\phi^2\right)\rho^2\phi \\
&\quad -\rho^2\phi^3V_{1,l}^2 \\
&\quad -V_{11}^2(1-\rho\phi)\phi - \rho\left(1-\phi^2\right)V_{11}^2 \\
&= (\phi+\rho)(1-\rho\phi)V_{11}V_y \\
&\quad -((\phi+\rho) + 2\rho^2\phi)(1-\rho\phi)V_{1,l}V_y \\
&\quad +(\rho+\phi)(1-\rho\phi)V_{1,l}V_y - \rho^2\phi V_{1,l}V_{11} \\
&\quad -\rho^2\phi^3V_{1,l}^2 \\
&\quad -(\phi+\rho)V_{11}^2 + 2\rho^2\phi V_{11}^2
\end{aligned}
$$

To obtain an approximate expression for α_1, we separate out the terms that are 'small' when both $|\rho|$ and $|\phi|$ are small (recall that both these parameters are small in our empirical model for real output growth):

$$
\begin{aligned}
&= (\phi+\rho)V_{11}V_y - (\phi+\rho)V_{1,l}V_y + (\phi+\rho)V_{1,l}V_{11} - (\phi+\rho)V_{11}^2 \\
&\quad [(\rho\phi^2 + \rho^2\phi)V_{11}V_y - 2\rho^2\phi(1-\rho\phi)V_{1,l}V_y - \rho^2\phi V_{1,l}V_{11} + 2\rho^2\phi V_{11}^2].
\end{aligned}
$$

The denominator (eq. 12) using 5B, 4B and 2B is:

$$
\begin{aligned}
& V(y_{t-1}^{t+l-1})V(y_{t-1}^{t}) - C(y_{t-1}^{t+l-1}, y_{t-1}^{t})^2 \\
= \; & [V_y (1-\rho\phi) - V_{1,l}(1-\rho\phi)] \times [V_y (1-\rho\phi) - V_{11}(1-\rho\phi)] \\
& - [V_y (1-\rho\phi) - V_{11} - \rho\phi V_{1,l}] \times [V_y (1-\rho\phi) - V_{11} - \rho\phi V_{1,l}] \\
= \; & -V_y V_{11} (1-\rho\phi)(1-\rho\phi) + 2V_{11} V_y (1-\rho\phi) \\
& - V_y V_{1,l} (1-\rho\phi)(1-\rho\phi) + 2V_y V_{1,l} (1-\rho\phi)\rho\phi \\
& + V_{11} V_{1,l} (1-\rho\phi)(1-\rho\phi) - 2\rho\phi V_{11} V_{1,l} \\
& - \rho^2\phi^2 V_{1,l} \\
& - V_{11}^2 \\
= \; & (1-\rho^2\phi^2) V_y V_{11} \\
& - (3\rho^2\phi^2 - 4\phi\rho + 1) V_y V_{1,l} \\
& + (\phi^2 - 4\phi\rho + 1) V_{11} V_{1,l} \\
& - \rho^2\phi^2 V_{1,l} \\
& - V_{11}^2
\end{aligned}
$$

Again separating out terms that are small when $|\rho|$ and $|\phi|$ are small:

$$= V_y V_{11} - V_y V_{1,l} + V_{11} V_{1,l} - V_{11}^2 + [\rho^2\phi^2 V_y V_{11} - (3\rho^2\phi^2 - 4\phi\rho) V_y V_{1,l} - \rho^2\phi^2 V_{1,l}]$$

And ignoring the terms in square brackets in the expressions for the numerator and denominator, we find:

$$\alpha_1 \approx \frac{(\phi+\rho)V_{11}V_y - (\phi+\rho)V_{1,l}V_y + (\phi+\rho)V_{1,l}V_{11} - (\phi+\rho)V_{11}^2}{V_y V_{11} - V_y V_{1,l} + V_{11}V_{1,l} - V_{11}^2} \approx \phi + \rho.$$

The population value of α_1 is not exactly equal to $\phi + \rho$ because we have ignored terms which are powers of ϕ and ρ or enter as products of these two parameters.

The numerator of α_2, using expressions 5B, 1B, 2B and 3B and ignoring noise terms, when $v_t^{t+l} \neq 0$ and $\phi \neq 0$ and revisions are news is:

$$
\begin{aligned}
= \;& \left[V_y\left(1-\rho\phi\right)-V_{1,l}\left(1-\rho\phi\right)\right]\times\left[V_y\rho\left(1-\rho\phi\right)-\rho\left(1-\phi^2\right)V_{11}-\rho\phi^2V_{1,l}\right] \\
& -\left[V_y\left(1-\rho\phi\right)-V_{11}-\rho\phi V_{1,l}\right]\times\left[V_y\rho\left(1-\rho\phi\right)+\phi V_{11}-(\rho+\phi)V_{1,l}\right] \\
= \;& -V_yV_{11}\left(1-\rho\phi\right)\rho\left(1-\phi^2\right)-V_yV_{11}\left(1-\rho\phi\right)\phi+V_{11}V_y\rho\left(1-\rho\phi\right) \\
& -V_yV_{1,l}\left(1-\rho\phi\right)\rho\phi^2-V_{1,l}V_y\rho\left(1-\rho\phi\right)\left(1-\rho\phi\right)+V_yV_{1,l}\left(1-\rho\phi\right)\left(\rho+\phi\right)+V_{1,l}V_y\rho^2\phi\left(1-\rho\phi\right) \\
& +V_{1,l}V_{11}\left(1-\rho\phi\right)\rho\left(1-\phi^2\right)-\left(\rho+\phi\right)V_{11}V_{1,l}+\rho\phi^2V_{1,l}V_{11} \\
& +V_{1,l}^2\left(1-\rho\phi\right)\rho\phi^2-\rho\phi\left(\rho+\phi\right)V_{1,l}^2 \\
& +\phi V_{11}^2 \\
= \;& V_yV_{11}\left(1-\rho\phi\right)\left(\rho\phi^2-\phi\right) \\
& +V_yV_{1,l}\left(1-\rho\phi\right)\left(\phi+2\rho^2\phi-\rho\phi^2\right) \\
& +V_{1,l}V_{11}(\rho^2\phi^3-\rho^2\phi-\phi) \\
& +V_{1,l}^2(\rho\phi^2-\rho^2\phi^3-\rho^2\phi-\rho\phi) \\
& +\phi V_{11}^2
\end{aligned}
$$

Separating out small terms as before:

$$
\begin{aligned}
= \;& -\phi V_yV_{11}+\phi V_yV_{1,l}-\phi V_{1,l}V_{11}+\phi V_{11}^2 \\
& [(2\rho\phi^2-\rho^2\phi^3)V_yV_{11}+(2\rho^2\phi-2\rho\phi^2-\rho^2\phi^2+\rho^2\phi^3)V_yV_{1,l}+ \\
& (\rho^2\phi^3-\rho^2\phi)V_{1,l}V_{11}+(\rho\phi^2-\rho^2\phi^3-\rho^2\phi-\rho\phi)V_{1,l}^2]
\end{aligned}
$$

gives rise to an approximation to the population value of α_2 of:

$$
\alpha_2 \approx \frac{-\phi V_yV_{11}+\phi V_yV_{1,l}-\phi V_{1,l}V_{11}+\phi V_{11}^2}{V_yV_{11}-V_yV_{1,l}+V_{11}V_{1,l}-V_{11}^2} \approx -\phi
$$

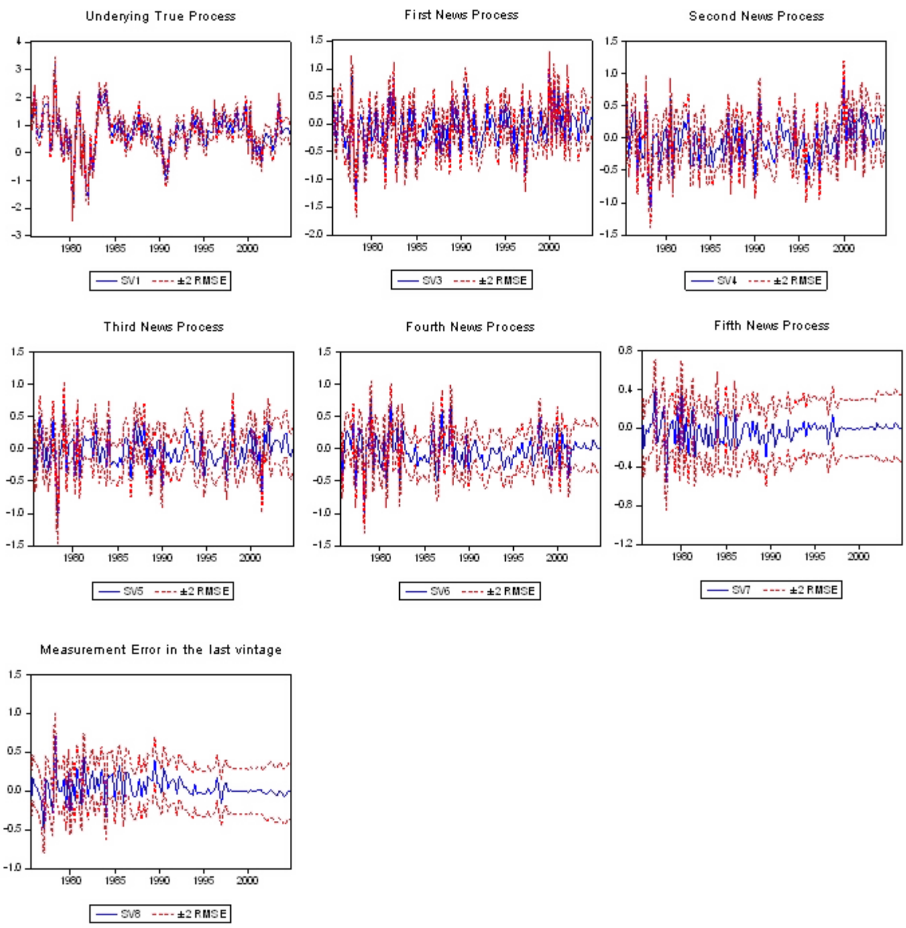

Figure 1: Smoothed estimates of the revisions processes from the state-space model.

Table 1: Regressions of final output on lagged final output and first announcements

Final output	2008Q2	$t+24$	$t+12$		$t+24$
Sample	1965:4 to 2001:4	1965:4 to 2001:4	1965:4 to 2001:4		1965:4 to 2001:4
Const.	0.537***	0.416***	0.389***	Const.	0.424***
	(.097)	(.105)	(.110)		(.105)
y_{t-1}	-0.031	0.135	0.269*	y_{t-1}	-0.178
	(.132)	(.185)	(.165)		(.320)
y_{t-1}^{t}	0.409***	0.290	0.175	y_{t-1}^{t+11}	0.572*
	(.133)	(.189)	(.183)		(.321)
σ	0.819	0.807	0.794	σ	0.805
R^2	0.132	0.159	0.181	R^2	0.163
Sample	1967:1 to 1991:4	1967:1 to 1991:4	1967:1 to 1991:4		1967:1 to 1991:4
Const.	0.541***	0.375***	0.367***	Const.	0.367***
	(.113)	(.121)	(.129)		(.123)
y_{t-1}	-0.101	0.143	0.218	y_{t-1}	-0.113
	(.154)	(.235)	(.204)		(.386)
y_{t-1}^{t}	0.465***	0.275	0.223	y_{t-1}^{t+11}	.508
	(.154)	(.232)	(.216)		(.385)
σ	0.912	0.905	0.898	σ	0.904
R^2	0.141	0.159	0.178	R^2	0.161
Sample	1974:3 to 2001:4	1974:3 to 2001:4	1974:3 to 2001:4		1974:3 to 2001:4
Const.	0.511***	0.418***	0.390***	Const.	0.416***
	(.101)	(0.110)	(.111)		(.110)
y_{t-1}	0.092	0.165	0.305*	y_{t-1}	-0.107
	(.147)	(.205)	(.170)		(.366)
y_{t-1}^{t}	0.289**	0.239	0.124	y_{t-1}^{t+11}	0.491
	(.146)	(.207)	(.191)		(.361)
σ	0.793	0.811	0.787	σ	0.809
R^2	0.134	0.146	0.175	R^2	0.151

Note. The dependent variable y_t is 'final' output, which is taken to be either $\left\{y_t^{2008Q2}\right\}$, $\left\{y_t^{t+24}\right\}$ or $\left\{y_t^{t+12}\right\}$. In the final column we replace the regressor y_{t-1}^{t} by y_{t-1}^{t+11}.
The figures in parentheses are robust (Newey-West) standard errors, σ is the estimated regression standard error, and ***, ** and * denote significance at the 1%, 5% and 10% levels respectively.

Table 2: Tests of the news and noise hypotheses

Revised Data	$2008:Q2$	$t+24$	$t+12$	$t+24$	$2008:Q2$
Initial Data	$t+1$	$t+1$	$t+1$	$t+12$	$t+24$
H_0: News	7.42[.001]	2.46[.089]	2.32[.101]	1.06[.348]	9.38[.000]
H_0: Noise	17.08[.000]	11.39[.000]	10.42[.000]	1.36[.324]	7.28[.000]

Note. Sample period is 1965:Q3-2001:Q4. The entries are the F-statistics (robust Wald statistic) of the null hypotheses that H_0: $\alpha = \beta = 0$ in regressions (7) and (8), followed by p-values in brackets.

Table 3: Monte Carlo estimates based on the estimated SSM

$T = 150$					
Simulated final data	$y^{08:Q2}$	y^{t+24}	y^{t+12}		y^{t+24}
y_{t-1}	0.245	0.293	0.293	y_{t-1}	0.293
	[.520]	[.584]	[.494]		[.216]
y_{t-1}^{t}	0.146	0.099	0.098	y_{t-1}^{t+11}	0.076
	[.187]	[.117]	[.109]		[.077]
$T = 1,000$					
Simulated final data	$y^{08:Q2}$	y^{t+24}	y^{t+12}		y^{t+24}
y_{t-1}	.251	.299	.298	y_{t-1}	.298
	[.999]	[1]	[.999]		[.828]
y_{t-1}^{t}	.156	.108	.106	y_{t-1}^{t+11}	.084
	[.744]	[.418]	[.354]		[.123]

Note. Number of replications is 10,000. Value in [] is the proportion of rejections of the null of no statistical significance at the 5% level using robust standard errors over the replications.
We simulate 50 additional observations in each, which are then discarded, to approximate a stationary draw.

Table 4: Estimates and Rejection Frequencies for the Rodriquez Mora and Schulstad (2007) regression based on the estimated VAR models

			$T = 150$				
VAR:	$q = 24$	$q = 12$	$q = 24$ News imp.	$q = 12$ News imp.		$q = 12$	$q = 12$ News imp.
Simulated Final Data:	y^{t+24}	y^{t+12}	y^{t+24}	y^{t+12}		y^{t+24}	y^{t+24}
y_{t-1}	0.065	0.162	-0.026	0.019	y_{t-1}	-.085	-.083
	[.114]	[.241]	[.099]	[.070]		[.102]	[.089]
y_{t-1}^{t}	0.221	0.126	0.267	0.242	y_{t-1}^{t+11}	.444	.435
	[.329]	[.152]	[.381]	[.347]		[.437]	[.395]
			$T = 1000$				
	$q = 24$	$q = 12$	$q = 24$ News imp.	$q = 12$ News imp.		$q = 12$	$q = 12$ News imp.
Simulated Final Data:	y^{t+24}	y^{t+12}	y^{t+24}	y^{t+12}		y^{t+24}	y^{t+24}
y_{t-1}	0.065	0.167	-0.029	0.022	y_{t-1}	-.082	-.079
	[.263]	[.851]	[.119]	[.076]		[.193]	[.164]
y_{t-1}^{t}	0.239	0.130	0.287	0.245	y_{t-1}^{t+11}	.449	.441
	[.987]	[.630]	[.997]	[.977]		[.997]	[.992]

Note. Number of replications is 10,000. Value in [] is the proportion of rejections of the null of no statistical significance at the 5% level with robust standard errors over the replications.
We simulate 300 additional observations in each replication, which are then discarded, to approximate a stationary draw.