

Nowcasting, Business Cycle Dating and the Interpretation of New Information when Real-Time Data are Available*

by

Kevin Lee,[†] Nilss Olekalns^{††} and Kalvinder Shields^{††}

Abstract

A canonical model is described which reflects the real-time informational context of decision-making. Comparisons are drawn with ‘conventional’ models that incorrectly omit market-informed insights on future macroeconomic conditions and inappropriately incorporate information that was not available at the time. It is argued that conventional models are misspecified and misinterpret news but that these deficiencies will not be clearly exposed either by diagnostic tests applied to the conventional models or by typical impulse response analyses. This is demonstrated through an analysis of quarterly US data 1968q4-2008q4. However, estimated real-time models considerably improve out-of-sample forecasting performance, provide more accurate ‘nowcasts’ of the current state of the macroeconomy and provide more timely indicators of the business cycle. The point is illustrated through an analysis of the US recessions of 1990q3–1991q2 and 2001q1–2001q4 and the most recent experiences of 2008.

Keywords: Structural Modelling, Real-Time Data, Nowcasting, Business Cycles.

JEL Classification: E52, E58

^{*†}University of Leicester, UK; ^{††}University of Melbourne, Australia. We have received helpful comments from participants at the Computational Economics and Finance Conference 2007, Canada, the Australasian Macroeconomics Workshop 2008, the Nowcasting and Model Combination Workshop, RBNZ, 2008 and seminar participants at the Reserve Bank of Australia. Version dated January 2009. Corresponding author: Kalvinder K. Shields, Department of Economics, University of Melbourne, Victoria, 3010, Australia. E-mail: k.shields@unimelb.edu.au, tel: 00 613-83445300, fax: 00 613-83446899.

1 Introduction

The availability of detailed real-time data sets, consisting of the successive vintages of data that have been released over time, makes it possible to systematically analyse the informational context in which decisions are made. It has been argued that this could be important both for understanding policy decisions made in the past and for providing policy advice in the future (see Blinder, 1997, Orphanides et al, 2000, Orphanides, 2001, and Cogley and Sargent, 2005 for example). However, the use of real-time data in macroeconomic analysis remains relatively rare. This paper considers the circumstances and extent to which models based on real-time data can improve on more conventional models focusing both on the interpretation of new information as it arrives and on decision-making that is sensitive to accurate business cycle dating.

Real-time data focuses attention on two interrelated aspects of the informational context within which macroeconomic decisions are made. The first is concerned with *end-of-sample issues* that arise because decisions are based on the currently-available data in the presence of measurement and future uncertainty. This aspect includes questions surrounding the relevant methods for forecasting (or "nowcasting" if information on today's position is published only with a delay) and the problems encountered in finding the appropriate model and econometric techniques to minimise and accommodate forecast uncertainties. These problems are compounded when variables are measured with error and there has developed a considerable literature concerned with the need to accommodate the effects of revisions in published data when making decisions at the end-of-sample (as highlighted in Kishor *et al* (2003) or Croushore and Evans (2006), for example).

The second aspect of macroeconomic decision-making highlighted by real-time data concerns *processing issues*. These include the interpretation of the "news" that becomes available in each period and has been discussed in the literature concerned with the identification of monetary policy or other types of shocks and the corresponding impulse response analyses. Here, attention has often focused on the timing and sequencing of decisions to obtain impulse responses that describe the impact of particular policy

innovations.¹ The role of the timing of decisions in the analysis highlights that it is important to use the relevant vintage of data to properly reflect the real-time informational context. But ‘processing issues’ also include questions on how to best interpret and exploit the wealth of information that is available to describe the current and expected future prospects of the economy, including direct measures of expectations available from surveys and market information (in financial markets, say).² The macroeconomic implications of the processing issues have been highlighted by Mankiw, Reis and Wolfers (2003), for example, who propose a ‘sticky-information’ model to explain business cycle properties. Similarly, Orphanides and Williams (2002) emphasise the importance of understanding the nature of inflation expectations to properly interpret the successes and failures of monetary policy. And, more recently, Gali and Gertler (2007) discuss the significance of the role of private sector expectations of the future performance of the economy and future policy actions in the monetary transmission mechanism.³

The importance of the two aspects of real-time data for any modelling activity will depend on the purpose of the modelling. If the purpose of the analysis is to test a particular economic theory or to interpret and understand past policy episodes, then the analysis is likely to place more emphasis on the interpretation and processing of data. If the purpose of the modelling is to facilitate real-time decision-making and forecasting, then the focus will be on end-of-sample issues. The aim of this paper, then, is to evaluate the use of real-time data in modelling in principle and in practice, distinguishing between the use of the data in interpreting new information and its use in nowcasting/forecasting and in making decisions influenced by business cycle conditions. In the first part of the paper, we describe a canonical model that explicitly reflects the real-time information

¹Important examples are provided in Bernanke and Mihov (1998), Sims and Zha (1998), Christiano *et al.* (1999), Rotemberg and Woodford (1999) and Brunner (2000), *inter alia*.

²There is a well-established empirical literature showing that direct measures of expectations contain useful information to explain future economic outcomes; see Batchelor (1986), Lee (1994), Smith and McAleer (1995), Roberts (1995, 1997) and Lee and Shields (2000) on the use of surveys, and Estrella and Trubin (2006) and Bordo and Haubrich (2008) on the use of yield curve information, *say*.

³The literature on learning is also relevant; see Evans and Honkapohja (2001) for a comprehensive review.

context of decision-making, accommodating the end-of-sample and the processing issues raised above. The model captures the simultaneous determination of first-release measures of macroeconomic variables, their expected future values and their subsequent revisions and illustrates the (typically extensive) restrictions required to identify economically-meaningful relations and the associated structural innovations. The canonical model also provides a means of considering the nature of conventional models found in the literature. Conventional models are based on “final vintage” datasets which measure variables after all the revisions have taken place and they omit the direct measures of expected future values that existed at the time. They both incorporate information on revisions that was not available at the time and ignore survey-based and market-informed insights on future macroeconomic conditions that were available. These provide the standard framework within which macroeconometric analysis takes place though, so it is useful to consider how the results of conventional analyses should be interpreted in the light of the canonical model framework and whether standard statistical tools will expose the misspecification.

The second part of the paper examines the extent to which these issues are important empirically through an analysis of quarterly US data over the period 1968q4 – 2008q4 using the empirical counterparts of the canonical and the conventional models. The analysis shows that the misspecification of the conventional models is not exposed either by diagnostic tests applied to the models or by typical impulse response analysis. However, their out-of-sample forecasting performance is considerably weaker than that of the fully-specified real-time model. The real-time analysis is particularly powerful in providing ‘nowcasts’ to accurately describe the current state of the macroeconomy and in providing more timely indicators of the business cycle. Therefore, in contrast to the conclusions of Croushore and Evans (2006), we argue that real-time considerations do have practical significance for economic decision-making and policy analysis. The power of the real-time model is illustrated through a case study of the use of information in the timely recognition of the recessions experienced in the US in 1990q3–1991q2 and 2001q1–2001q4 and of the most recent experiences of 2008.

The layout of the paper is as follows. Section 2 introduces the canonical modelling framework proposed to take into account the information available in real time. The

section illustrates the identification issues involved and discusses the links between this model and the conventional models typically found in the literature. Section 3 describes the US real-time data, including a summary of the sequencing of data releases. It also introduces three alternative and increasingly sophisticated empirical models that can be estimated making progressively greater use of the data available in real time. This aims to establish quantitatively the importance of taking into account the various sources of information available in real time in macroeconomic policy analysis. Section 4 reports on the use of the models in constructing nowcasts and forecasts of recessions in real time. The section describes a case study analysing the information flows that would have informed decision makers in the recession of 2001q1 – 2001q4 and compares this with the use of information in the recession that occurred a decade earlier in 1990q3 – 1991q3. The use of information in recognising and dating the onset of the current recession is also investigated. Section 5 concludes.

2 A Modelling Framework to Accommodate Real-Time Information

In this section, we describe a canonical macroeconomic model that is able to accommodate explicitly the information available in real time, including the release of different vintages of data and of direct measures of expected future outcomes. Comparison of the canonical model with a “conventional” model of macroeconomic dynamics helps establish the ways in which real-time data might improve macroeconomic analysis and those where it might be less useful. Broadly-speaking, we argue that the real-time data are important in dealing with ‘end-of-sample’ issues but are probably less helpful in addressing ‘processing issues’.

In what follows, ${}_t x_{t-s}$ is the measure of the (logarithm of the) variable x at time $t - s$ as released at time t and ${}_t x_{t+s}^e$ is a direct measure of the expected value of the variable at $t + s$, with the expectation formed on the basis of information available at the time the measure is released, t . The sample of data runs from $t = 1, \dots, T$. We write ${}_t \mathbf{x}_t = ({}_t x_{1t}, \dots, {}_t x_{mt})'$, an $m \times 1$ vector of variables, so that \mathbf{x}_t might be a 4×1 vector containing data on the interest rate, output growth, price inflation, and money growth, for example. For ease of exposition, we assume in the first instance that the determination and first-measurement of variables is synchronised and that data are revised once following its first

release. In this case, the model is:

$$\mathbf{A}_{11} \mathbf{x}_t = -\mathbf{A}_{12} \mathbf{x}_{t+1}^e - \mathbf{A}_{13} \mathbf{x}_{t-1} + \mathbf{B}_{11} \mathbf{x}_{t-1} + \mathbf{B}_{12} \mathbf{x}_t^e + \mathbf{B}_{13} \mathbf{x}_{t-2} + \boldsymbol{\varepsilon}_{bt} \quad (2.1)$$

$$\mathbf{A}_{22} \mathbf{x}_{t+1}^e = -\mathbf{A}_{21} \mathbf{x}_t - \mathbf{A}_{23} \mathbf{x}_{t-1} + \mathbf{B}_{21} \mathbf{x}_{t-1} + \mathbf{B}_{22} \mathbf{x}_t^e + \mathbf{B}_{23} \mathbf{x}_{t-2} + \boldsymbol{\varepsilon}_{et} \quad (2.2)$$

$$\mathbf{A}_{33} \mathbf{x}_{t-1} = -\mathbf{A}_{31} \mathbf{x}_t - \mathbf{A}_{32} \mathbf{x}_{t+1}^e + \mathbf{B}_{31} \mathbf{x}_{t-1} + \mathbf{B}_{32} \mathbf{x}_t^e + \mathbf{B}_{33} \mathbf{x}_{t-2} + \boldsymbol{\varepsilon}_{rt} \quad (2.3)$$

where \mathbf{A}_{ij} and \mathbf{B}_{ij} , $i, j = 1, 2, 3$, are $m \times m$ matrices of coefficients and $\boldsymbol{\varepsilon}_{bt}$, $\boldsymbol{\varepsilon}_{et}$ and $\boldsymbol{\varepsilon}_{rt}$ are $m \times 1$ vectors of shocks with mean zero and diagonal covariance matrices Ω_b , Ω_e and Ω_r , respectively. We can normalise the diagonal elements of \mathbf{A}_{11} , \mathbf{A}_{22} and \mathbf{A}_{33} to unity so that the equations of the system explain, respectively, the time- t measure of each of the variables in \mathbf{x}_t , the time- t expectation of \mathbf{x}_{t+1} and the time- t revised measures of \mathbf{x}_{t-1} .⁴ The structural model (2.1)-(2.3) reflects the fact that three interrelated processes occur here simultaneously and in real time: (i) ‘behavioural’ economic decisions are made by economic agents to determine the actual values of the variables at each time; (ii) expectations are formed on the variables by those same economic agents; and (iii) the economic outcomes are measured reflecting the data collection and survey practices of the statistical agencies.

The equations in (2.1)-(2.3) can be stacked to obtain

$$\mathbf{A} \mathbf{z}_t = \mathbf{B} \mathbf{z}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (2.4)$$

where $\mathbf{z}_t = (\mathbf{x}_t, \mathbf{x}_{t+1}^e, \mathbf{x}_{t-1})'$, $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{23} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} \end{bmatrix}$ and

$\boldsymbol{\varepsilon}_t = (\boldsymbol{\varepsilon}_{bt}, \boldsymbol{\varepsilon}_{et}, \boldsymbol{\varepsilon}_{rt})'$ with covariance $\Omega = \begin{bmatrix} \Omega_b & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Omega_e & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Omega_r \end{bmatrix}$. The corresponding reduced

form VAR is

$$\mathbf{z}_t = \mathbf{C} \mathbf{z}_{t-1} + \mathbf{u}_t, \quad (2.5)$$

⁴The equation in (2.3) can obviously be written in the ‘revision’ form $\mathbf{A}_{33} (\mathbf{x}_{t-1} - \mathbf{x}_{t-1}) = \dots + (\mathbf{B}_{31} - \mathbf{A}_{33}) \mathbf{x}_{t-1} + \dots$.

where $\mathbf{C} = \mathbf{A}^{-1}\mathbf{B}$ and $\mathbf{u}_t = \mathbf{A}^{-1}\boldsymbol{\varepsilon}_t$ with covariance matrix $\Sigma = \mathbf{A}^{-1}\Omega\mathbf{A}^{-1}$. Given that the contemporaneous interactions between variables are accommodated explicitly in the \mathbf{A} matrix, it is typically assumed that the structural innovations in Ω are orthogonal to each other.⁵ Identification of the parameters of the structural model in (2.4), and the associated structural innovations, from the parameters in (2.5) requires $\frac{9m^2+3m}{2}$ restrictions based on *a priori* theory, although subsets of the parameters and innovations might be identified on the basis of fewer relevant restrictions.

It is worth reflecting on how the canonical model of (2.4) and the associated reduced form of (2.5) relate to the more conventional models of macroeconomic variables found in the literature. These ignore real-time considerations by making use only of the final vintage of data and, usually, by eschewing the direct measures of expectations that are available. The relationship between the models becomes clearer by noting that conventional models effectively focus on the post-revision series ${}_t\mathbf{x}_{t-1}$ only. If data are revised only once (and remains unchanged thereafter), then ${}_T\mathbf{x}_{t-1} = {}_t\mathbf{x}_{t-1}$, $t = 1, \dots, T$. In this case, apart from the observation at the end of the sample, the final vintage of data ${}_T\mathbf{X}_t$ is the same series as the post-revision series ${}_t\mathbf{X}_{t-1}$ measured at time T :

$${}_T\mathbf{X}_t = \{ {}_T\mathbf{x}_1, {}_T\mathbf{x}_2, \dots, {}_T\mathbf{x}_{T-2}, {}_T\mathbf{x}_{T-1}, {}_T\mathbf{x}_T \},$$

$${}_t\mathbf{X}_{t-1} \text{ available at time } T = \{ {}_2\mathbf{x}_1, {}_3\mathbf{x}_2, \dots, {}_{T-1}\mathbf{x}_{T-2}, {}_T\mathbf{x}_{T-1}, \emptyset \},$$

where \emptyset represents a missing entry. If the sample of data is sufficiently long (so that the difference at the end of the series is unimportant), any model estimated using the final vintage of data will be equivalent to that obtained using the post-revision series only. Further, the canonical model of \mathbf{z}_t determines the nature of the model that should be estimated for any subset of the variables in \mathbf{z}_t . The reduced form (2.5) means $\mathbf{z}_t = (\mathbf{I} - \mathbf{CL})^{-1}\mathbf{u}_t$ and the time series model of any variable or subset of variables in \mathbf{z}_t is determined by the lag structure of $(\mathbf{I} - \mathbf{CL})^{-1}$ and the properties of the \mathbf{u}_t . The point is illustrated simply if we consider \mathbf{x}_t to contain just one variable. In this case, the system

⁵Realistically even structural innovations could be correlated with each other. The orthogonality assumption implicitly means that one of the innovations is reconfigured so that any correlation is subsumed into the structural parameters in \mathbf{A} and the innovation is redefined as that part of the original which is orthogonal to the rest.

in (2.5) is a three-variable VAR of order 1. But each individual series also admits a univariate ARMA(3,2) specification so that we can write, for example,

$${}_t x_{t-1} = \lambda_1 {}_{t-1} x_{t-2} + \lambda_2 {}_{t-2} x_{t-3} + \lambda_3 {}_{t-3} x_{t-4} + v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2}, \quad t = 1, \dots, T, \quad (2.6)$$

where the λ_1 , λ_2 and λ_3 are functions of the parameters in \mathbf{C} while the θ_1 , θ_2 and properties of the errors v_t are determined by matching its correlelogram with that of the combination of shocks given by $(\mathbf{I} - \mathbf{CL})^{-1}\mathbf{u}_t$.⁶ Estimation of (2.6), or the corresponding univariate autoregressive approximation, will provide the same estimates of the λ and θ parameters whether we use the post-revision series or the final vintage series, subject to the sample not being dominated by the differences at the end of the data.

2.1 Real-time data and information processing

An argument put forward for the use of models that employ real-time data is that they can help clarify the processing issues surrounding the use of information and the interpretation of news. The canonical model of (2.4) illustrate the potential for identification provided by the detail of real-time data. But it also shows that considerable *a priori* information is required to define meaningful structural relations, and this will include information on the nature and measurement of data and on the expectation formation process as well as on the decision-making of economic agents. For example, if the system of (2.4) included data on the four variables, interest rates, output, inflation and money, we would require 78 identifying restrictions to be imposed to identify the structural model and all of the underlying behavioural shocks from the associated reduced form VAR. The availability of these restrictions will vary from one modelling exercise to another, but it is worth elaborating on three particular sources here to clarify the real-time issues that are involved.

One source of potential identifying restrictions is from economic theory. For example, recently, there have been many ‘New-Keynesian’ models described in the literature

⁶Noting that $(\mathbf{I} - \mathbf{CL})^{-1} = [\det(\mathbf{I} - \mathbf{CL})]^{-1}\text{adj}(\mathbf{I} - \mathbf{CL})$, the autoregressive element is based on $\det(\mathbf{I} - \mathbf{CL}) = 1 - \lambda_1 L - \lambda_2 L^2 - \lambda_3 L^3$ and the moving average element depends on the combination of reduced form errors given $\text{adj}(\mathbf{I} - \mathbf{CL})$. See Hamilton (1994, p. 349) for details.

which have clearly specified micro-foundations and which provide well-defined dynamic relationships between key macroeconomic variables; see the references in Gali and Gertler (2007), for example. If measurement issues are ignored, then such models can be readily accommodated within a VAR framework and the structure suggested by the theory provides (many) over-identifying restrictions, on the contemporaneous and lagged parameters, with which the theory can be tested.⁷ The identifying restrictions for recovering these structural relationships in estimation would be more complex, however, if the analysis takes into account the nature of the information set that is available in real time. This is because the microfoundations of a New Keynesian model would require assumptions to be made not only on firms' price setting behaviour but also on *which* information the firms use to form expectations; i.e. whether they based their decisions on first-releases of published data or on expectations of post-revision data, and so on.⁸ Agents in the model would also need to form a view on the extent to which the statistical agency publishes the 'raw' data obtained as the outcome of a clearly defined data collection exercise (even if this includes systematic measurement error of unknown source) or whether the agency attempts to purge the data of systematic error prior to publication.⁹ In short, the demands on the economic theory used for identification are much more challenging in the presence of real-time data.

A second potential source of identifying restrictions on the system in (2.4) is through assumptions on the nature of the expectations formation process. The characterisation of expectations in (2.1) makes no assumptions on the expectation formation process but the model can accommodate many alternative assumptions through the imposition of restrictions on its parameters. This includes the rational expectation (RE) hypothesis, for example. However, implementing the identifying restrictions arising from the RE assumption in (2.4) requires assumptions to be made on which measure of the variables agents had in mind when reporting their expectations. For example, the identifying

⁷See Kim and Pagan (1995) or Pesaran and Smith (2006), for example.

⁸See Croushore and Evans's (2006) related discussion on whether policy decisions are based on observed first release data or the 'true' underlying state of the economy when estimating policy rules.

⁹See Jacobs and van Norden (2006) for discussion of the sources of revision error in published data and the extent to which the revisions reflect the 'news' or 'noise' described in Mankiw and Shapiro (1983).

restrictions relating to (2.2) will be quite different depending on whether respondents in a survey report their expectation of the first release measure, so ${}_{t-1}\mathbf{x}_t^e = E[{}_t\mathbf{x}_t \mid I_{t-1}]$, or they report their expectation of the “actual” post-revision measure, so ${}_{t-1}\mathbf{x}_t^e = E[{}_{t+1}\mathbf{x}_t \mid I_{t-1}]$. In the former case, for example, we can write $\mathbf{B}_{12} = \mathbf{I}_m$, $\mathbf{A}_{12} = \mathbf{A}_{13} = \mathbf{B}_{12} = \mathbf{B}_{13} = \mathbf{0}$ so that ${}_t\mathbf{x}_t = {}_{t-1}\mathbf{x}_t^e + \boldsymbol{\varepsilon}_{et}$ in (2.1). The identifying structure would be quite different if ${}_{t-1}\mathbf{x}_t^e = E[{}_{t+1}\mathbf{x}_t \mid I_{t-1}]$ however and, again, the presence of real-time data complicates the identification based on assumptions on the expectation formation process.

A third form of *a priori* information used to provide identifying restrictions on the system in (2.4) is through assumptions on the timing and/or sequencing of decisions. These are particularly common in studies of the effects of monetary policy. The assumptions are typically used to motivate a diagonal, or block diagonal, structure in the ‘contemporaneous’ matrix corresponding to \mathbf{A} in VAR models of macroeconomic variables. In this case, the use of real-time data sometimes allows a more accurate description of the timing of data releases and decisions and this might help some, if not all, shocks to be identified. For example, given that interest rate data are available at any point during a quarter while quarterly output, price and money data are published at various specified times during the quarter, one could choose to use the beginning-of-quarter measure of the interest rate to unambiguously place this *first* in the sequence of behavioural decisions determining these four variables.¹⁰ On the further assumption that interest rate and other forecasts are also determined after the policy decision, then we can place the interest rate first in the vector of variables \mathbf{z}_t and write the first row of $\mathbf{A} = (1, 0, 0, 0, \dots)'$ in (2.4). The reduced form equation for the interest rate will then provide an estimate of the structural interest rate equation and the shocks to the reduced form interest rate equation will have the standard interpretation as reflecting structural monetary policy shocks.¹¹

This discussion shows that the use of real-time data might help in identifying and processing information in some circumstances, especially where the sequencing of deci-

¹⁰A more detailed description of the timing of US data releases is provided in the following section.

¹¹Garratt et al. (2006) illustrate that, assuming that interest rates are set ‘first’, a structural interest rate equation of this form can be derived as the outcome of the optimising decisions of a monetary authority faced with a structural model of the form in (2.4) and an objective function that is quadratic in some or all of the other variables in \mathbf{x}_t .

sions is well understood. But the discussion also makes clear that the extra detail of the data typically requires a corresponding increase in the detail of the structure provided by *a priori* information for it to be useful in identification. In the absence of this detailed *a priori* structure, models that employ real-time data are unlikely to have any advantage over simpler conventional models in providing insights on behavioural/structural interpretations of news, testing economic theory, policy evaluation, and so on. Further, the real-time models can highlight difficulties in interpretation which are obscured in more conventional models. As a specific example, the discussion surrounding (2.6) showed that, even if the real-time model of (2.4) reflects the true data-generating process, there will be a perfectly admissible time series representation of the post-revision series taken in isolation from the other data and this representation can usually be estimated in a conventional model using the final-vintage data series only. However, (2.6) also showed that the shocks in this conventional model are actually a convoluted combination of the true structural innovations.¹² The real-time data highlights that the straightforward interpretation of impulse responses based on conventional models should be treated with caution therefore.

2.2 Real-time data and diagnostic testing

The discussion above suggests that, in the absence of detailed *a priori* information on the structural relations, the real-time and conventional models are equally admissible as representations of the macroeconomy. In these circumstances, we might evaluate the use of the real-time data by testing the statistical adequacy of the models to reflect the data. However, as the discussion surrounding the simplified model at (2.6) made clear, the conventional models will typically provide an entirely adequate representation of the data even if the true data generating process is as in the real-time model of (2.4). The conventional model will have a more complicated ARMA structure than the original and, in practice, these might be approximated with high-order AR models in estimation. Standard diagnostic tests (on serial correlation, functional form, non-normality, outliers,

¹²As (2.6) makes clear, the time- t innovation in the lower-dimension model consists of shocks to expected future values of the series and to revisions as well as shocks to first-release and revised measures of the series.

and so on) might appear poor if the approximation is poor. But in principle, if the real-time model is well-specified, then there is no reason to expect the diagnostic tests for the conventional model to indicate misspecification.

An alternative approach to testing the statistical adequacy of conventional models in the presence of data revisions is by looking for instability in models estimated using different data vintages; see Orphanides (2001) or Croushore and Evans (2006), for example. However, the arguments presented above suggest that if the underlying real-time model is an adequate representation of the data, then the corresponding conventional model will also be an adequate representation. If the sample size is large enough for the final observation to have an insignificant impact on the estimation, the exercise will provide sensible, unbiased estimates of the parameters of the conventional model (which are functions of the true underlying parameters of (2.4)). These estimates will remain stable over time, irrespective of the vintage of data used in estimation.¹³

The structural stability of the conventional model is based on the assumption that the real-time model is itself stable over time. Structural instability in the real-time model will, of course, translate to structural instability in the conventional model. The extent to which this misspecification shows in diagnostic tests of the conventional model will depend on the details of the real-time model. But it is intuitively reasonable to think that the conventional model might be relatively sensitive to structural breaks because a structural break in any one of the relationships of the real-time model will show as instability in the corresponding conventional model. Generally speaking, then, diagnostic testing of the conventional model is unlikely to expose model misspecification, although structural instability tests might be more sensitive than other tests.

¹³The univariate conventional model of (2.6) corresponded to a three-variable VAR of order 1 where there is one revision only. If the data contains p revisions, then the multivariate model would be a $(p + 2)$ -variable VAR and the corresponding univariate representation of the post-revision series would be ARMA($p + 2, p + 1$). Nevertheless, the same arguments would apply.

2.3 Real-time data and forecasting

The most striking deficiency of conventional models relative to those using real-time data is likely to be exposed in nowcasting and forecasting exercises. Assume we are interested in a straightforward nowcast of the current state of the business cycle at the end-of-sample period T . We have assumed so far that the determination and first measurement of the variables are synchronised so the first-release observation on the data at the end-of-sample Ty_T is available automatically in the time- T vintage of data. However, this will provide a biased estimate of the post-revision nowcast measure in which we are interested if there is a predictable element in revisions. The measurement error will contaminate subsequent forecasts at longer horizons and the characterisation of the uncertainty surrounding the forecasts will also be incorrect. Forecast statements on more complicated events (e.g. forecasting the probability of two consecutive periods of negative growth, for example, or another business cycle feature) will be unreliable for the same reasons.¹⁴

In contrast, the estimated real-time model of (2.5) can be used in a straightforward way to produce the one-step-ahead forecast of the post-revision nowcast measure $T+1y_T$ alongside the forecasts of the next period's survey expectation $T+1y_{T+2}$ and the first-release measure of $T+1y_{T+1}$. The forecast performance, at the one-step ahead and subsequent forecast horizons, will depend on the context and it is widely recognised that forecasts based on estimated versions of the true data generating process can be outperformed by simpler misspecified models, in a mean squared error sense, if the true model includes variables with relatively little explanatory power (see Clements and Hendry, 2005). However, the direct measures of expectations included in the real-time model are likely to have good explanatory power; indeed, the direct measures will themselves provide the optimal forecast if expectations are formed rationally and relate to the post-revision measure. Further, there is considerable evidence that systematic and predictable elements exist in data revisions and these will also contribute to a real-time model's forecasting performance.

In brief, it seems likely that conventional models which do not make use of direct measures of expectations or take into account revisions in data will be particularly poor at

¹⁴See Garratt *et al* (2003) for discussion of the simulation methods that underlie the production of event probability forecasts.

forecasts that focus on the end of the sample when decisions are made. This is ultimately an empirical issue, however, and we therefore explore the relative forecast performance of various models of US macro data in the following section.

3 The Informational Content of US Real-Time Data

In this section, we provide an analysis of US data on output growth, inflation, money and interest rates to investigate the information content of the first-releases of measures of the series, of revisions in the data and of direct measures of expectations of the variables. The real-time dataset is obtained from the Federal Reserve Bank of Philadelphia at www.phil.frb.org/econ/forecast/ and consists of 172 quarterly vintages of data; the first was released in 1965q4 and the final vintage used in this paper is dated 2008q4. All vintages include variable observations dated back to 1947q1. The analysis in this section is primarily statistical aimed at illustrating the issues raised in the previous section to evaluate the usefulness of real-time data. The usefulness is judged first in the context of identifying and tracing out the macroeconomic effects of monetary policy shocks and then in the context of ‘nowcasting’ and forecasting the current and future state of the macroeconomy.

3.1 Timing of US Data Release

The empirical analysis starts with a description of our macroeconomic data taking proper account of the timing of the data releases in the US. For aggregate output, data on real GDP in quarter t is released for the first time at the end of the first month of quarter $t + 1$. This figure is reported in the Federal Reserve Bank of Philadelphia’s real-time data set as the mid-point of the $(t + 1)^{th}$ quarter and is denoted by ${}_{t+1}y_t$ in what follows, where y_t is the logarithm of real GDP. Revisions that subsequently take place in output measures in the months up to the mid-point of the $(t + 2)^{nd}$ quarter are reported in ${}_{t+2}y_t$. Likewise, ${}_{t+3}y_t$ incorporates any revisions that are then made up to the mid-point of the $(t + 3)^{th}$ quarter, and so on.

Money and price measures are released monthly with a one month delay. In this

analysis, p_{t-1} refers to the *average* value of the (logarithm of) the consumer price index (CPI) over the three months of quarter $t - 1$. The observation for prices in the third month of quarter $t - 1$ is not released until the end of the first month of quarter t and so, matching the timing of the release of the output data, we take each quarter's price observation to be released at the mid-point of the succeeding quarter, denoted ${}_t p_{t-1}$. So, for example, the average data for the months that constitute the first quarter, January, February and March, are assumed to become available in the following May; the average data for the months that constitute the second quarter, April, May and June, are assumed to become available in the following August, and so on. The timing of the release of data on the M1 measure of the money supply is exactly the same and so ${}_t m_{t-1}$ also refers to the average of the data relating to the three months of quarter $t - 1$ released for the first time at the mid-point of quarter t .

Our measure of the rate of interest, r_t , is the Federal Funds rate. The Federal Reserve's Open Market Committee usually meets eight times a year; in February, March, May, July, August, September, November and December and the outcome of its deliberations are immediately made known. The decision on how to measure the rate at the quarterly frequency is relatively arbitrary, and so we can choose to measure the rate in a way that justifies any assumptions on the timing of interest rate decisions. To be consistent with the assumption that interest rate decisions are made first within the quarter, we take as our measure of the quarterly interest rate, ${}_t r_t$, the Federal Funds rate as observed at the beginning of January, April, July, and October, i.e. the interest rate holding on the first day of the relevant quarter.

To investigate the informational content of 'forward-looking' variables, we make use of the interest rate spread (to reflect market expectations of future rates) and experts' forecasts on output and prices as provided in the Federal Reserve Bank of Philadelphia's *Survey of Professional Forecasters* (SPF), from 1968q4 – 2008q4. The spread is denoted ${}_t sp_t$ and is defined as the difference between the three-month Treasury Bill Secondary Market Rate, converted to a bond-equivalent basis, and the market yield on US Treasury securities at a 10 year constant maturity (quoted on investment basis).¹⁵ Both series are

¹⁵See Estrella and Trubin (2006) for discussion.

obtained from the *H.15: Selected Interest Rates* publication of the Board of Governors of the Federal Reserve System. The observations for the spread are taken at the beginning of each quarter to coincide with the interest rate series. Forecasts taken from the SPF are made around the mid-point of quarter t although, in fact, the forecasters have available to them the first release information on the previous quarter's output and price level, ${}_{t-1}y_{t-1}$ and ${}_{t-1}p_{t-1}$ at the time when the forecasts are made.¹⁶ The nowcasts relating to quarter t 's output and price level are denoted by ${}_{t-1}y_t^f$ and ${}_{t-1}p_t^f$, and the forecasts of quarter $t + s$ output and price level, $s > 0$, are denoted by ${}_{t-1}y_{t+s}^f$ and ${}_{t-1}p_{t+s}^f$, respectively.

3.2 Model Specifications

To investigate the informational content of the various data that become available, we estimate three simple macroeconomic models which make increasingly specialised use of the data: a 'conventional' model which ignores real-time considerations; a specification that pays attention to the timing of data releases and revisions but does not include any forward-looking information; and a model which includes all the information available in real time. Following the discussion of the previous section, the intention is to produce a VAR analysis of interest rates, output, prices and money. Preliminary investigation shows that interest rates are stationary but output, prices and money series are integrated of order one and need to be differenced to obtain stationarity. In our models, we consider output growth, price inflation and money growth in the analysis, measuring these using changes in the (log) of the first-release data. As shown in Garratt et al (2006), a model that explains this growth measure alongside the revisions data is entirely justifiable statistically on the assumption that growth in the respective series is stationary and that measurement errors and expectational errors are all stationary.¹⁷ In all three models, shocks to interest

¹⁶Given that the spread information becomes available at the start of quarter t , the SPF will have internalised this source of contemporaneous information also.

¹⁷Garrat *et al.* (2006) show that the VAR model in first-release growth, expected growth and revisions can be written as a cointegrating VAR. This equivalent representation incorporates cointegrating relations between the levels of the first-release measure, expected value and revised value of each variable with cointegrating vectors $(1, -1, 0)$ and $(1, 0, -1)$. In the empirical part of the paper, we find that these restrictions cannot be rejected. Working with this system precludes the possibility of any long-run

rates and growth rates die out in the infinite horizon but have persistent effects on the levels of output, prices and money.

Model 1; Specification with Conventional Timing The first model we consider is a simple four-variable Vector Autoregressive Model explaining interest rates, output growth, price inflation and money growth using the final vintage data series only; i.e. a model of the form in (2.5), using

$$\mathbf{z}_t = ({}_T r_t, ({}_T y_t - {}_T y_{t-1}), ({}_T p_t - {}_T p_{t-1}), ({}_T m_t - {}_T m_{t-1}))',$$

for $t = 1, \dots, T - 1$. The timing of this model is ‘conventional’ in the sense that this is the form of the data that is typically employed in macroeconomic analysis. Here, the investigator considers only the most recent (time- T) data series available, assuming that these were the data available at the time decisions were made (presumably subject to some innocuous measurement error) and effectively ignoring the fact that revisions have taken place. Further, the data here are aligned temporally on the basis of the time period t to which the observation refers, not of the date of release. This assumes that all of the data that relate to time period t were available at time period t despite the publication delays known to operate in practice. This model provides the baseline comparator, therefore, abstracting from all real-time considerations.

Model 2; Specification with Real-Time Data and Revisions Our second model specification takes into account the release of information at each point in time, estimating a model of the form in (2.5), using

$$\mathbf{z}_t = ({}_t r_t, ({}_t y_{t-1} - {}_{t-1} y_{t-2}), ({}_t p_{t-1} - {}_t p_{t-2}), ({}_t m_{t-1} - {}_t m_{t-2}), \\ ({}_t y_{t-2} - {}_{t-1} y_{t-2}), ({}_t y_{t-3} - {}_{t-1} y_{t-3}))',$$

for $t = 1, \dots, T$. This model includes the real-time measures of the four macroeconomic series of interest, measured taking into account the one-quarter publication lag described earlier, plus two output revisions. The model more realistically replicates the decision-making context faced by agents using information actually known to policy makers and relations between the levels of the interest rate, output, prices and money however.

other economic agents at the time at which decisions are made. Simple variable exclusion tests lead us to include up to two revisions of output in the model only and to drop revisions in money and prices altogether.¹⁸

Model 3; Specification with Real-Time Data, Revisions, and Economic Indicators Our third model specification supplements the system of Model 2 with direct measures of expectations of current and future economic activity available in real time, estimating a model of the form in (2.5), using

$$\mathbf{z}_t = \left(\begin{array}{l} {}_t r_t, \quad ({}_t y_{t-1} - {}_{t-1} y_{t-2}), \quad ({}_t p_{t-1} - {}_t p_{t-2}), \quad ({}_t m_{t-1} - {}_t m_{t-2}), \\ ({}_t p_t^f - {}_t p_{t-1}), \quad ({}_t y_t^f - {}_t y_{t-1}), \quad ({}_t p_{t+1}^f - {}_t p_t^f), \quad ({}_t y_{t+1}^f - {}_t y_t^f), \quad {}_t s p_t \\ ({}_t y_{t-2} - {}_{t-1} y_{t-2}), \quad ({}_t y_{t-3} - {}_{t-1} y_{t-3}) \end{array} \right)',$$

for $t = 1, \dots, T$. The model therefore includes, in addition to the variables of Model 2, time- t measures of the nowcast of inflation and output growth from the SPF, direct measures of one-quarter ahead forecasts of the same series and the interest rate spread.

3.3 Estimation Results and Impulse Response Functions

A real-time analysis of the models will involve their recursive estimation at each point in time.¹⁹ However, useful insights on the nature of the conventional analyses of Model 1 can be obtained, and compared to the real-time analyses of Models 2 and 3, by looking in detail at examples of the estimated models based on a particular sample. Tables 1 and 2 therefore report the estimated VARs of Models 1-3 based on the final vintage of data for Model 1 (so $T = 2008q4$, $t = 1967q1, \dots, T - 1$) and on the real time data for Models 2 and 3 (so $T = 2008q4$ and $t = 1968q4, \dots, T$).

¹⁸To be more specific, although the money and price series are revised, these revisions have no systematic, statistically significant pattern and their lagged values make no significant contribution to the explanation of the other variables in the system. Similar comments apply to the third (and longer) revisions in output. Test results are available from the authors on request.

¹⁹For Model 1, this might involve a recursive analysis of the final vintage data, using the appropriate sample periods but using measures of the data which would not have been available at the time. This is termed “quasi real-time analysis” by Orphanides and van Norden (2002).

Model 1 The results show that there is considerable complexity in the feedbacks between the variables, with standard variable addition tests showing that a VAR of order 4 is appropriate (although lagged money appears to have a relatively minor role in explaining interest rates, growth or inflation). Simple inspection of the estimated coefficients indicates that strong growth and/or high inflation precede interest rate rises, as might be expected with a “Taylor-type” rule, that interest rate rises are associated with a subsequent slowdown in growth, and that inflation is influenced by positive growth with a long (four quarter) lag.

This overview is confirmed by the impulse response functions (IRFs) plotted in Figure 1, which show the impact of a shock to the interest rate equation to each of the four variables. This is typically interpreted as a monetary policy shock on the assumption that interest rates are set ‘first’, as discussed earlier. The IRFs show the effect of a monetary policy shock that raises interest rates by one standard error on impact with the rate returning to the level obtained in the absence of the shock after one or two years. The output response is quite protracted, with relatively strong effects lasting some two-three years, including a substantial fall in output relative to the base for over a year (so that output levels will be approximately 0.1% lower at the infinite horizon than in the absence of the shock). The inflation response reflects the ‘price puzzle’ often featured in the literature, whereby the interest rate rise is associated with a *rise* in inflation on impact and a small negative/neutral effect in the long run. And the response of money is a substantial reduction in money holdings, both in the short and long term. Stated briefly, then, the ‘conventional’ system equations appear complex but sensible in terms of the signs and magnitudes of the coefficients and the overall system properties are exactly of the form typically found in empirical exercises of this kind.

The diagnostic statistics in Table 1 also suggest that the four equations in this specification are reasonable ones according to the fit and, generally speaking, to the absence of evidence of serial correlation, functional form problems, heteroscedasticity or non-normality in the residuals. The main indicator of problems with the model is the strong evidence of structural instability, at least in the interest rate, inflation and money equations, identi-

fied through the application of the standard F-test to the sample split in half at 1986q1.²⁰ Taken at face value, then, Model 1 appears to provide a reasonable characterisation of the data and one that is broadly in line with macroeconomic stylised facts. However, there is evidence of instability which would render the model inappropriate for real-time forecasting or policy prescription even if the data were measured without error so that the final vintage data used here had been available at the time.

Model 2 Table 2 reports on Models 2 and 3 obtained using the first-release data and revisions in the series for $t = 1968q4, \dots, 2008q4$. The body of the table describes the estimated VAR for Model 2. This confirms that the analysis of data available in real time, including data on revisions, provides a distinct and even more complicated dynamic characterisation of the macroeconomic data than Model 1. Importantly, there are very clear, statistically-significant, systematic patterns in the first and second revisions of output, and the revisions themselves also play an important role in explaining the evolution of the (first-release measures of) output growth. The interest rate remains positively related to output growth and inflation and the signs of the short-run and long-run elasticities in the growth and inflation equations again appear sensible. But the size and the timing of the effects are quite different to those in Table 1, with this model able to accommodate the interrelatedness of measured output growth, its revision and their impact on the other macroeconomic variables which Model 1 cannot.

The coefficient estimates of Table 2 show clearly the statistical significance of separately modelling the first-release and revised measures of output. However, the differences between the models are obscured when considering the system-wide response to an interest rate shock. This is illustrated in Figure 1 where the effects of an interest rate shock on Model 2 are traced against those in Model 1. The interest rate is assumed to be set ‘first’ in both Models 1 and 2, so the shock has the same interpretation in both sets of impulses. Further, the impulses have been calculated to trace the effect of the shock on comparable output, inflation and money series in both models; namely, the *post-revision* output,

²⁰Subsequent tests suggest that there was a degree of stability during the first half of the sample (between 1967q1-1986q1) but evidence of further instability within the latter half.

inflation and money series (i.e. to ${}_{t+3+s}y_{t+s}$ $s = 0, 1, \dots$, in the case of output, where there are systematic revisions for two periods, and ${}_{t+1+s}p_{t+s}$ and ${}_{t+1+s}m_{t+s}$ for prices and money where the revisions have no systematic content). These series are approximately equal to the final vintage series used in Model 1, therefore.²¹ There are some differences between these responses (with the monetary policy shock having a more protracted effect on interest rates and output than in Model 1, for example). But it is perhaps surprising to find the impulse responses looking so similar in Models 1 and 2 given the statistical significance of the additional dynamics made explicit in Model 2. On reflection, however, this may not be so hard to understand. Specifically, we have already noted that, even if the VAR Model 2 is the true data generating process, it is possible to estimate a VARMA time series model for any sub-set of the variables in Model 2 which will approximate the true DGP. Having recognised that the post-revision series in Model 2 are approximately equal to the final-vintage series used in Model 1, it is clear that Model 1 can be interpreted as a simplified approximate version of Model 2. The estimated impulse responses of the post-revision series in Model 2 illustrate the same properties of the system dynamics captured by the responses of Models 1 to the same interest rate shock, therefore. This is reassuring if this particular impulse response exercise is the purpose of the analysis. But it is misleading if the model was to be used to trace the effect of other types of shock or in providing a structural interpretation to the estimated model.²²

²¹An impulse response function illustrates the time profile of a variable in response to a particular shock relative to the profile when no shock occurs. The definition of the responses of post-revision output to a shock specified by $\mathbf{u}_t = \bar{\mathbf{u}}$ for Model 1 is given by

$$E[{}_T y_{t+s} \mid I_{t-1}, \mathbf{u}_t = \bar{\mathbf{u}}] - E[{}_T y_{t+s} \mid I_{t-1}], \quad s = 1, \dots,$$

while the response of post-revision output to a shock specified by $\mathbf{u}_t = \underline{\mathbf{u}}$ for Model 2 is given by

$$E[{}_{t+3} y_{t+s} \mid I_{t-1}, \mathbf{u}_t = \underline{\mathbf{u}}] - E[{}_{t+3} y_{t+s} \mid I_{t-1}], \quad s = 1, \dots .$$

For impulse responses from different models to be comparable, the responses must relate to the impact of the same shock (so $\bar{\mathbf{u}} = \underline{\mathbf{u}}$).

²²The argument suggests that the two models would generate similar impulse responses of the post-revision series if the shock is the same in the two models. However, no shock can be specified that is defined similarly in both Models 1 and 2 apart from that to the interest rate.

The equation diagnostics again provide broad reassurance on the statistical coherence of the model according to fit and the standard residual-based tests. The evidence for structural instability is weaker for all the equations now with no evidence of instability working strictly at the 5% level of significance. In brief, then, the estimated equations of Model 2 also appear sensible in terms of signs and magnitudes of coefficients and have reasonable diagnostic properties.

Model 3 The lower section of Table 2 summarises the impact of adding to Model 2 the forward-looking variables suggested in Model 3, again focusing on the model estimated over $t = 1968q4, \dots, 2008q4$. A specification search suggested that six lags of the spread, ${}_tsp_t$, two lags of each of the SPF nowcasts ${}_ty_t^f$ and ${}_tp_t^f$ and two lags of the one-quarter ahead forecasts, ${}_ty_{t+1}^f$ and ${}_tp_{t+1}^f$ should be included in the equations and the $\chi_{LM}^2(14)$ statistic indicates the significance of these variables in each equation. The other three χ_{LM}^2 statistics isolate in turn the separate contributions of the spread, the SPF nowcasts, and the one-quarter ahead SPF forecasts. These confirm that all three series have considerable explanatory power in the interest rate, output growth and inflation equations. This highlights the potential misspecification problems of macroeconomic modelling exercises that omit forward-looking variables. Interestingly, the forward-looking data, and especially the spread, also provide significant explanatory power for the revisions series, suggesting that these data may reflect agents' expectation of the true underlying data.

The underlying short-run and long-run elasticities of Model 3 are not reported in Table 2 for space considerations. But they are sensible according to sign and magnitudes once more and provide reasonable system dynamics. Indeed, the impulse responses of the post-revision series to an interest rate shock based on Model 3 are again reported in Figure 1 and again correspond closely to those of the previous models (especially Model 1). However, Model 3 provides the most comprehensive description of the DGP for these macroeconomic series and the specifications of Models 1 and 2 can be interpreted as approximations that adequately capture the system dynamics (at least as far as these particular impulse responses are concerned) but would be misleading for more structural analysis.

The fit and diagnostic tests of Model 3 (not reported for space considerations but available on request) again show an improvement over the other models. Indeed, as the figures in the final row of Table 2 demonstrate, the inclusion of the additional forward-looking variables serves to eliminate any remaining evidence of structural instability. This is in itself an important empirical finding, showing that a VAR model that attempts to implicitly capture the effect of expectations formation in macroeconomic models is unlikely to succeed.²³ Model 3 represents our preferred model, therefore, accommodating directly all of the information that is available to decision-makers at the time decisions are made. But the weaknesses of Models 1 and 2 and the relative superiority of Model 3 were not at all obvious from the diagnostic test results and impulse responses of the first two models.

3.4 Model Evaluation using Statistical Forecasts

This section provides an evaluation of the out-of-sample point forecasting performance of the different models. The analysis focuses on forecasts of output growth and inflation at various horizons to judge the extent to which the use of the data on revisions and measures of expectations make a useful contribution if decisions are made in real time based on nowcasts or forecasts of these variables. Since forecasts from Model 1 are not directly comparable to those of Models 2 and 3, we estimated Model 1', a four-variable VAR of the form in (2.5) obtained in real time using

$$\mathbf{z}_t = \left(\begin{array}{c} {}_t r_t, \quad ({}_t y_{t-1} \quad {}_{t-1} y_{t-2}), \quad ({}_t p_{t-1} \quad {}_{t-1} p_{t-2}), \quad ({}_t m_{t-1} \quad {}_{t-1} m_{t-2}) \end{array} \right)'. \quad (3.7)$$

Table 3 reports root mean squared errors (RMSE's) for Models 1', 2 and 3, where the models are estimated recursively for $t = 1968q4, \dots, \tau$, and the relevant out-of-sample forecasts are computed at each recursion for up to two years ahead; i.e. at $\tau+h$, $h = 1, \dots, 8$. We chose $\tau = 1985q4, \dots, 2008q4 - h$ so that the RMSE's are based on up to $N = 80$ recursions. Four RMSE's are obtained using forecasts relating to output growth alone

²³This should not be interpreted as evidence against rationality in expectation formation. Rather, it suggests that the information content of the direct measures of expectations cannot be captured here by linear function of lagged values of the variables and simple structural shocks of Models 1 and 2.

and two are obtained relating to price inflation forecasts alone. Specifically, these are based on:

1. the nowcast of the first-release output level, $\widehat{\tau+1y_\tau} = E[\tau+1y_\tau | I_\tau]$, which effectively involves a one-step ahead forecast since output is released with a one quarter delay;
2. the nowcast of actual, post-revision output level, $\widehat{\tau+3y_\tau}$, which will involve three-quarter ahead forecasts accounting for the one-quarter delay in the release of output and for two quarterly revisions;
3. the forecast of actual output two-quarters ahead $\widehat{\tau+5y_{\tau+2}}$;
4. the forecast of actual output four-quarters ahead $\widehat{\tau+7y_{\tau+4}}$;
5. the nowcast of the first-release price series, $\widehat{\tau+1p_\tau}$, and
6. the forecast of prices four-quarters ahead $\widehat{\tau+4p_{\tau+3}}$.

In addition, we also report RMSE's based on functions of output and inflation forecasts that might be of more direct interest to decision-makers. Specifically, we also focus on

7. the nowcast of the output gap, $x_t | \Omega_{t+s} = {}_{t+3}y_t - \tilde{y}_t$, defined as the gap between actual output at t and the trend measure, \tilde{y}_t , obtained by running the Hodrick-Prescott filter through the forecast-augmented actual output series $\{\dots, {}_{t-1}y_{t-4}, {}_t y_{t-3}, \widehat{{}_{t+1}y_{t-2}}, \widehat{{}_{t+2}y_{t-1}}, \widehat{{}_{t+3}y_t}, \widehat{{}_{t+4}y_{t+1}}, \dots\}$.²⁴
8. the nowcast of a policy objective, $g_t | \Omega_t = \lambda(x_t | \Omega_t) + ({}_{t+1}p_t - {}_t p_{t-1})^2$ defined as a weighted aggregate of the output gap and inflation where the weight on the gap is varied from $\lambda = 0.1, 0.3, 0.5$.

The table also shows the outcome of two sets of tests of forecast accuracy. The first set is provided by the Diebold-Mariano (DM) statistics which test the null of equal predictive accuracy of Models 1' and 2 and then Models 2 and 3 respectively, based on the differences

²⁴Details of the computation of the gap measure are given in Garratt et al (2008), where the gap is based on a forecast-augmented Hodrick-Prescott smoother.

in the reported root mean square errors and an estimate of the asymptotic variance of this difference (see Diebold and Mariano (1995)). The second set of tests compares Models 2 and 3 only and is obtained from a simulation exercise based on the assumption that the estimated Model 2 obtained using data for $t = 1968q4, \dots, \tau$ is the true data generating process for $t = 1968q4, \dots, \tau + h$. Under this assumption, 100,000 replications of the data sample were generated. Then, for each replication r : Model $2^{(r)}$ and Model $3^{(r)}$ were estimated; forecasts were made for the period $\tau + 1, \dots, \tau + h$ using each model; corresponding $\text{RMSE}^{(r)}$ were calculated for the two alternative models; and the difference between these two RMSE's was recorded. The 100,000 simulated difference statistics obtained in this way provide an empirical distribution for the statistic under the null that Model 2 is true. The \dagger and \ddagger indicate whether the difference in RMSEs observed in the table is greater than the upper 10% or 5% of that empirical distribution. This test statistic is likely to be a more powerful test of the usefulness of the extra variables in Model 3 for forecasting than the DM test when comparing forecasts of nested models (see Clark and McCracken (2001)) and can be readily applied even when the prediction criterion is a complicated function of forecasts of different variables and over different forecast horizons.

Comparison of the RMSE statistics for Models 1' and 2 shows clearly that the revisions data are useful in the nowcasts of first-release and actual output growth. The RMSE of the output growth nowcasts from Model 2 are some 25% lower than those from Model 1' and the DM tests show this to be very strong evidence of improved forecast accuracy. The performance of the longer horizon forecasts of output growth, or for inflation, is not enhanced by the inclusion of the revision data (with the RMSE of Model 2 actually being worse, although not significantly so). This is not so surprising for the inflation series, where revisions were seen to be unimportant. But it also means that the improved forecasting performance achieved through inclusion of the revisions data is achieved primarily on nowcasts and is less pronounced for forecasting over the medium or longer term. This is not to deny its importance; the end-of-sample forecasting performance is crucial in real-time decision-making, for example. But it shows clearly where the gains arise.

Comparison of the RMSE for Models 2 and 3 show even more strikingly the usefulness in forecasting of including all the information available at the time decisions are made,

including direct measures and market-based measures of expectations. The RMSE errors calculated using Model 3 are substantially and significantly less than those calculated using Model 2 for all the forecasts considered, covering all the variables and combinations of variables at every horizon. Improvements of up to 40% in the RMSE are observed across the various criteria with the expectations data providing particular forecast improvement on the inflation series. It is worth emphasising that these results are found without using a very sophisticated specification search; we have noted the diagnostics used to choose appropriate lag lengths, for example, but there has been no further search conducted and many variables remain in the model with relatively low t-values. The clarity of the findings on the improved forecasting performance is not the outcome of sophisticated data-mining therefore but simply reflects the importance of including these explanatory variables and fully exploiting the information that is available to forecasters at the time forecasts and decisions are made.

4 The Usefulness of Real-Time Data for Nowcasting and Forecasting

The results of the previous section show that, in terms of purely statistical criteria, there is a strong argument for using real-time data, including direct and market-based expectations measures, in modelling. In this section, we show that the use of the available information in modelling and forecasting is equally important when judged using more economic criteria in the context of decision-making. To this end, we propose specific economic events of interest relating to the business cycle and use these as a basis for evaluating Models 1-3 by comparing the models' performance in forecasting the likelihood of the events taking place.

The calculation of probability forecasts (i.e. forecasts of the probability of specified events taking place) is relatively unusual in economics. This is surprising given that, compared to the point forecasts and confidence intervals that are usually reported, probability forecasts are better able to focus on events of interest to decision-makers and can convey the uncertainties associated with the event of interest more directly. Further, the methods are relatively straightforward to implement using simulation methods. Garratt et al. (2003) describe the methods in detail, but the idea can be briefly outlined if we

consider an example where we calculate the probability density function (pdf) associated with the nowcast of actual output growth defined by the change in post-revision output $(\widehat{y_{t+3}} - \widehat{y_{t+2}})$.²⁵ Here, one would use the estimates from a chosen model, including the estimated variance-covariance of the innovations, to generate R replications of the future outcomes by simulation. Each of the simulated futures includes values for $\widehat{y_{t-3+h}}^{(r)}$, $h = 0, 1, \dots$ where the ‘ (r) ’ superscript again denotes the value taken in the r^{th} simulation ($r = 1, \dots, R$). The values of $\widehat{y_{t-3+h}}^{(r)}$ obtained across replications directly provides the simulated pdf of forecast post-revision output at time $t - 3 + h$ and the values of $(\widehat{y_{t+3}}^{(r)} - \widehat{y_{t+2}}^{(r)})$ provide the pdf of the nowcast of actual output growth.²⁶ Further, counting the number of times in which $(\widehat{y_{t+3}}^{(r)} - \widehat{y_{t+2}}^{(r)})$ exceeds zero out of the R replications provides a direct estimate of the nowcast probability that output growth is positive. This statistic will be much more useful to a decision-maker concerned with this specific feature of the business cycle than the point forecast of growth and 95% confidence intervals typically reported.

To illustrate the importance of using real-time information in this context, we focus on two events relating to the time- t perception of the business cycle at time t . The first considers the likely occurrence of two periods of consecutive negative growth at t and $t - 1$ as measured by the ‘actual’ post-revision data; i.e. $\Pr\{A\}$ where event A is defined by $A : \{ [(y_{t+3} - y_{t+2}) < 0] \cap [(y_{t+2} - y_{t+1}) < 0] \}$. This is one simple but frequently used definition of “recession”. Figure 2 plots these forecast probabilities for $t = 1986\text{q}1 - 2008\text{q}4$ as calculated from the estimates of Model 2 (dashed line) and the estimates of Model 3 (solid line) obtained recursively in real time and on the basis of $R = 200,000$ replications. The figure shows that the event actually occurred in only two out of the 89 quarters of our observed sample upto 2008q1 (namely 1991q1 and 2001q4). Through the shading, it also reflects the fact that the relevant post-revision data are not available to judge whether the event has occurred during 2008q2-q4 at the time of writing

²⁵We abstract from parameter uncertainty in this example although this feature can be readily accommodated; see Garratt et al (2003) for details.

²⁶It is worth emphasising that this growth nowcast involves forecasts of series at different forecast horizons which are not independent. However, the simulated pdf automatically reflects all the uncertainties associated with these forecasts.

in 2008q4.

The nowcasts of the probability of event A occurring remain close to zero for both models in most periods, rising above 20% on just one occasion for Model 2 and on four occasions for Model 3 prior to 2008q1. Both models recognise the increased likelihood of recession in 1991q1, with the probability rising to 40% for Model 2 and 98% for Model 3. Strikingly, though, only Model 3 nowcasted a high probability of recession in 2001q4 in real time, providing a 79% probability of recession compared to Model 2's 6%. Similar disparities arise between the most recent probability forecasts from the two models and especially between those for 2008q4. Of course, these disparities partly reflect Model 3's ability to accommodate agents' real-time perception of the cycle in addition to the backward-looking data underlying the forecasts of Model 2. In 2008q4, the remarkable events of September 2008 had taken place²⁷ and negative growth seemed assured: the SPF nowcasted growth at an annual rate of -2.9% in 2008q4, for example (this compares to the forecast of +0.7% for the quarter made by the SPF in 2008q3 and a point-nowcast of a similar order of magnitude from Model 2). This new information obviously has a considerable impact on the nowcast probability of recession based on Model 3 compared to Model 2. But, importantly, this extra information also influences the forecast of the size and direction of revision of the first-release measure of growth *in the previous period*. The first-release measure of growth in 2008q3 was close to zero (an annualised rate of -0.3%) but this translates to a forecast probability that the actual growth was negative of 0.69 for Model 2 compared to 0.89 for Model 3. The bad news of 2008q4 causes an evaluation of the released data for 2008q3 which shifts the likely timing of the onset of recession defined in this way to 2008q3. Time will tell whether Model 3's relative pessimism was justified in 2008q4 as it was in 2001q4.

²⁷The third quarter of 2008 marked several significant economic events for the US economy which prompted widespread anxiety about global recession. The global investment bank Lehman Brothers Holdings Inc. filed for bankruptcy on September 15th marking the largest bankruptcy in US history. This led to a loss in confidence in inter-bank lending and caused widespread panic in global financial markets. September also saw bailouts by the US government of mortgage companies Fannie Mae and Freddie Mac (a \$100 billion bailout), of the insurance giant AIG (\$150 billion bailout), and a credit freeze for many businesses as banks hoarded cash for bad mortgages and withdrawals in bank runs.

The second business cycle event considered here is the occurrence of recession as defined by the NBER (see www.nber.org). The NBER definition of recession is based on a number of economic indicators and the recession dates are published only after a significant delay. For instance, the end of the recession in 2001q4 was only announced by the NBER in July 2003. In our exercise, we evaluate our alternative models from the perspective of decision-makers who need to know whether we are in an NBER-defined recession today. The first step in this process is to relate the NBER categorisation to observable data. To this end, a probit model is estimated to explain a dummy variable, $NBER_t$, which takes a value of one for all quarterly dates of contraction as defined by the NBER and zero otherwise. Following a relatively straightforward specification search, eliminating longer lags when statistically insignificant, the regressors in the model consist of the current and one lag of actual output growth (${}_{t+3}y_t - {}_{t+2}y_{t-1}$), and the current and one lag of a dummy variable based on the ‘current depth of recession’. This is defined as the gap between the current level of actual output and its historical maximum; i.e. $\max\{{}_{t+3}y_{t-s}\}_{s=0}^{\infty} - {}_{t+3}y_t$. The related dummy variable, denoted CDR_t , will take the value of one when output dips below its previous maximum and zero otherwise.²⁸ The estimated Probit model obtained using data for 1965q4 – 2007q4 is as follows:

$$\begin{aligned}
 NBER_t = & \underset{[-1.2594]}{-0.6452} - \underset{[-3.4663]}{158.3721}({}_{t+3}y_t - {}_{t+2}y_{t-1}) - \underset{[-2.0243]}{58.8598}({}_{t+2}y_{t-1} - {}_{t+1}y_{t-2}) \\
 & + \underset{[0.0579]}{0.0316}CDR_t + \underset{[2.3515]}{1.1532}CDR_{t-1} + \hat{\epsilon}_t,
 \end{aligned}$$

where $\epsilon_t \sim N(0, 1)$ and where t-statistics are reported in [.]

To calculate the “nowcast” probabilities of an NBER-recession, it is assumed that the relationship between $NBER_t$ and the measurables in (4.8) is fixed throughout our sample. We then calculate $\Pr\{B\}$ where event B is defined by $B : \{ NBER_t > 0 \}$, obtained recursively in real time using the same the $R = 200,000$ simulations of the future as described above. The nowcast values of $NBER_t$ are complicated non-linear functions of forecasts of output measures at different forecast horizons, so the likely occurrence of an NBER-recession would be extremely difficult to calculate analytically. The estimated

²⁸The asymmetry implied by the CDR term is reflected in the “bounce-back” effect, the tendency for output growth to recover relatively strongly following a recent recession.

probabilities are relatively easily obtained through the simulation exercise, however, and are illustrated for 1986q1-2008q4 in Figure 3. This figure shows that contraction was actually observed, according to the NBER, in ten of the quarters considered in the diagram; namely during 1990q3–1991q3 and 2001q1–2001q4 and, as announced in December 2008, in 2008q1. NBER pronouncements on the three subsequent quarters is unavailable at the time of writing and this region is depicted by the shaded area. Model 2 performs relatively poorly in identifying these periods in real time. The nowcast probability of NBER-contraction based on Model 2 exceeds 50% on only two occasions through the period and only one corresponds to periods subsequently labelled as contractions by NBER. Model 3 on the other hand performs relatively well, with the nowcast probability exceeding 50% on nine occasions, seven of which correspond to NBER dates.²⁹ The correspondence between the probability forecasts based on Model 3 and the event outcomes is striking in Figure 3. This shows again the considerable information content of survey data and market-based expectations in judging where the economy currently stands.

As a final exercise here, Table 4 provides further details of the nowcasted probabilities for the periods of contraction identified by NBER to uncover the nature of the information content contained in the survey and yield curve data. The first row of the Table shows the probabilities reported in Figure 3 and based on Model 3 including the spread data plus the current realisations and one-quarter ahead expectations of inflation and output growth. The subsequent three rows show the corresponding probabilities obtained if only the spread data were included in the model, only the realisation data were included, and only the one-step ahead expectations data were included, respectively. The results in these three rows are based on misspecified models (having incorrectly dropped statistically significant variables) and should be treated with caution. But they provide indicative information on the source of the information useful in forecasting. As it turns out, the relatively high probabilities (>35%) observed in 1990q3–1991q2, 2001q1–2001q4

²⁹The forecasts of Model 3 have placed a high likelihood that the NBER will identify the period as a contraction through 2008. There is, however, a dip in the probability plot in 2008q3 to 30%. This reflects the relative optimism of forecasters on growth shown as late as August 2008 when the 2008q3 surveys are conducted and when, for example, the SPF nowcasted a annualised output growth of 1.2%. As noted above, the events of September caused analysts to revise their opinions rapidly in 2008q4.

and 2008q1 in Model 3 appear to be driven primarily by the use of the survey-based realisation data. The one-step ahead expectations data are useful too, if used in isolation, but it is the realisation data which allows the model to rapidly identify the state of the business cycle. The lower half of the table reports in an analogous fashion the contraction probabilities for the periods but based on information available one year before the contraction. Interestingly, if we focus on the figures that show reasonably high ($>20\%$) contraction-probabilities, this set of results indicates that it is the spread data which is most useful at this longer one-year ahead forecast horizon.

5 Concluding Comments

Real-time datasets are becoming increasingly available. The literature on real-time analysis contains both an enthusiasm and hope that these datasets can deliver new insights on macroeconomic policy-making and macroeconomic research and also a scepticism on the usefulness of the results obtained so far (see Croushore's (2008) review). This paper aims to pinpoint the areas in which real-time data is likely to be important, both from a theoretical and empirical perspective.

The discussion surrounding the model of Section 2 and its relationship with more conventional models concluded that real-time data is likely to be important in dealing with end-of-sample issues surrounding forecasting and exercises concerned with dating the business cycle. However, real time data would be less helpful in addressing the processing issues involved in the identification of structural models. The empirical work confirmed that this is true in the case of the US since 1968. A real-time model that uses data available at the time, including survey- and market-based information, has satisfactory diagnostics and sensible system dynamics. The model appears superior to simpler real-time and conventional models by these criteria but the misspecification of the latter models was not self-evident from their own diagnostic test statistics or system properties. The use of real-time data highlights the difficulties of structural modelling using conventional models but the real-time model does not provide any easy solutions to the processing problems themselves.

On the other hand, there are substantial gains from the use of real-time data in

the practice of nowcasting and forecasting, business-cycle dating and decision-making. Measured by a range of statistical criteria, the performance of the real-time model in revealing the current and future business cycle position was shown to be substantially and significantly better than the simpler models. This was true whether the analysis focused on specific variables (output or prices) or more complicated functions (involving gaps or other policy objectives). The result was found at all forecast horizons but was particularly strong for contemporaneous nowcasts. This conclusion, based on statistical criteria, was reinforced by the performance of the real-time model in producing density forecasts and event probability forecasts relating to the US recessions of the early 1990's, the early 2000's and most recently. A full evaluation of these event forecasts requires a full description of the loss function faced by the decision-maker. But the ability of the real-time model to reveal periods of recession as they occurred was striking. This is the case whether recession is considered as two consecutive periods of negative growth or the more complicated conjuncture of events considered by NBER. The evidence that decision-makers can gain timely insights from real-time analysis goes some way to justifying the enthusiasm shown by some empirical economists for the analysis of real-time datasets.

Table 1: Model 1: VAR with Conventional Timing: 1967q1 - 2008q3

Independent Variable	Dependent Variable			
	Tr_t	$(Ty_t - T y_{t-1})$	$(Tp_t - T p_{t-1})$	$(Tm_t - T m_{t-1})$
intercept	-0.0151 (0.0049)	0.0065 (0.0020)	-0.0073 (0.0010)	0.0033 (0.0023)
Tr_{t-1}	0.2633 (0.0912)	-0.0958 (0.0373)	-0.0212 (0.0194)	0.0468 (0.0428)
Tr_{t-2}	0.1196 (0.0938)	0.0293 (0.0373)	0.0114 (0.0199)	-0.0512 (0.0440)
Tr_{t-3}	0.3150 (0.0936)	0.0847 (0.0372)	-0.0310 (0.0199)	-0.0029 (0.0439)
Tr_{t-4}	0.1600 (0.0950)	-0.0415 (0.0378)	0.0201 (0.0202)	0.0307 (0.0446)
$(Ty_{t-1} - T y_{t-2})$	0.4784 (0.2062)	0.1637 (0.0820)	0.0718 (0.0438)	-0.2200 (0.0968)
$(Ty_{t-2} - T y_{t-3})$	0.5257 (0.2088)	0.2086 (0.0831)	0.0104 (0.0443)	-0.0541 (0.0980)
$(Ty_{t-3} - T y_{t-4})$	0.3570 (0.2129)	0.0103 (0.0847)	0.0293 (0.0452)	0.0673 (0.0999)
$(Ty_{t-4} - T y_{t-5})$	0.0742 (0.1888)	0.0240 (0.0751)	0.1128 (0.0401)	0.0672 (0.0886)
$(Tp_{t-1} - T p_{t-2})$	1.0871 (0.3964)	-0.1864 (0.1577)	0.6257 (0.0842)	-0.4303 (0.1861)
$(Tp_{t-2} - T p_{t-3})$	-0.0438 (0.4179)	0.0092 (0.1663)	-0.0203 (0.0887)	0.3954 (0.1962)
$(Tp_{t-3} - T p_{t-4})$	0.5545 (0.4229)	-0.0644 (0.1682)	0.5453 (0.0898)	-0.2807 (0.1958)
$(Tp_{t-4} - T p_{t-5})$	-0.3395 (0.4136)	0.0792 (0.1645)	-0.1566 (0.0878)	0.2729 (0.1942)
$(Tm_{t-1} - T m_{t-2})$	0.0477 (0.1892)	0.0606 (0.0753)	0.0278 (0.0402)	0.5366 (0.0888)
$(Tm_{t-2} - T m_{t-3})$	-0.0768 (0.2152)	0.1324 (0.0856)	0.0318 (0.0457)	0.2068 (0.1010)
$(Tm_{t-3} - T m_{t-4})$	-0.1410 (0.2144)	-0.0833 (0.0853)	0.0477 (0.0455)	0.0648 (0.1006)
$(Tm_{t-4} - T m_{t-5})$	0.1230 (0.1856)	-0.0095 (0.0738)	-0.0194 (0.0394)	-0.0600 (0.0871)
R^2	0.7835	0.2695	0.7495	0.5265
$\hat{\sigma}$	0.0183	0.0073	0.0039	0.0086
$F_{SC(4)}$	{0.11}	{0.51}	{0.28}	{0.01}
F_{FF}	{0.19}	{0.03}	{0.00}	{0.37}
F_H	{0.00}	{0.35}	{0.14}	{0.28}
F_N	{0.00}	{0.00}	{0.14}	{0.03}
F_{STAB}	{0.01}	{0.51}	{0.00}	{0.00}

Notes: Standard errors are given in (.). R^2 is the squared multiple correlation coefficient, and $\hat{\sigma}$ is the standard error of the regression. The remaining diagnostics are p-values, in { . }, for F-test statistics for serial correlation (SC), functional form (FF), normality (N), heteroscedasticity (H), and a Chow test of the stability of regression coefficients (STAB).

Table 2: Model 2: VAR with real-time Data and Revisions: 1968q4 - 2008q4

Independent Variable	Dependent Variable					
	$t^r t$	$(t y_{t-1} -_{t-1} y_{t-2})$	$(t p_{t-1} -_{t-1} p_{t-2})$	$(t m_{t-1} -_{t-1} m_{t-2})$	$(t y_{t-2} -_{t-1} y_{t-2})$	$(t y_{t-3} -_{t-1} y_{t-3})$
intercept	-0.0159 (0.0051)	0.0044 (0.0020)	0.0015 (0.0011)	0.0018 (0.0022)	-0.0021 (0.0010)	-0.0016 (0.0100)
$t^r t-1$	0.3316 (0.0850)	-0.0371 (0.0333)	0.0782 (0.0186)	-0.1638 (0.0360)	-0.0370 (0.0174)	-0.0134 (0.0151)
$t^r t-2$	0.0253 (0.0913)	-0.0508 (0.0358)	-0.0360 (0.0200)	0.1010 (0.0387)	0.0484 (0.0186)	0.0324 (0.0163)
$t^r t-3$	0.4003 (0.0966)	-0.0082 (0.0379)	-0.0058 (0.0212)	-0.0029 (0.0410)	-0.0126 (0.0197)	-0.0101 (0.0172)
$t^r t-4$	0.0970 (0.0913)	0.0617 (0.0358)	-0.0322 (0.0200)	0.0640 (0.0387)	-0.0026 (0.0186)	-0.0039 (0.0163)
$(t-1 y_{t-2} -_{t-2} y_{t-3})$	0.8909 (0.2536)	0.5244 (0.0994)	0.0600 (0.0555)	-0.0797 (0.1075)	0.1433 (0.0518)	0.0502 (0.0452)
$(t-2 y_{t-3} -_{t-3} y_{t-4})$	0.1727 (0.2547)	0.0919 (0.0998)	-0.0719 (0.0558)	0.1232 (0.1080)	0.0289 (0.0520)	0.0472 (0.0454)
$(t-1 p_{t-2} -_{t-2} p_{t-3})$	0.3881 (0.3657)	-0.4242 (0.1433)	0.5943 (0.0801)	-0.2133 (0.1550)	-0.0662 (0.0746)	-0.0589 (0.0651)
$(t-2 p_{t-3} -_{t-3} p_{t-4})$	0.7472 (0.3826)	0.3468 (0.1499)	0.2054 (0.0837)	0.3525 (0.1622)	0.1896 (0.0781)	0.0974 (0.0682)
$(t-1 m_{t-2} -_{t-2} m_{t-3})$	0.0454 (0.1890)	0.0108 (0.0741)	0.0572 (0.0413)	0.5564 (0.0801)	0.0589 (0.0386)	0.0348 (0.0337)
$(t-2 m_{t-3} -_{t-3} m_{t-4})$	0.0069 (0.1883)	0.1211 (0.0738)	-0.0092 (0.0412)	0.2105 (0.0800)	-0.0100 (0.0384)	0.0131 (0.0335)
$(t-1 y_{t-3} -_{t-2} y_{t-3})$	-1.1962 (0.8955)	-1.0262 (0.3579)	0.1326 (0.1961)	-0.8524 (0.3800)	-0.5370 (0.1823)	-0.2971 (0.1595)
$(t-2 y_{t-4} -_{t-3} y_{t-4})$	1.3278 (0.9134)	0.5592 (0.3579)	0.0121 (0.2000)	-0.3354 (0.3872)	-0.2074 (0.1864)	-0.2505 (0.1627)
$(t-1 y_{t-4} -_{t-2} y_{t-4})$	0.2024 (0.9709)	0.7283 (0.3804)	-0.2915 (0.2126)	1.0127 (0.4115)	0.3506 (0.1982)	0.2101 (0.1730)
$(t-2 y_{t-5} -_{t-3} y_{t-5})$	-0.9969 (0.9720)	-0.6629 (0.3809)	0.0736 (0.2129)	0.1845 (0.4120)	0.1935 (0.1984)	0.1846 (0.1732)
R^2	0.7741	0.3918	0.7164	0.5900	0.1757	0.1013
$\hat{\sigma}$	0.0188	0.0073	0.0074	0.0118	0.0040	0.0033
$F_{SC(4)}$	{0.25}	{0.87}	{0.00}	{0.04}	{0.13}	{0.24}
F_{FF}	{0.94}	{0.13}	{0.02}	{0.05}	{0.83}	{0.63}
F_H	{0.00}	{0.68}	{0.36}	{0.00}	{0.90}	{0.36}
F_N	{0.00}	{0.27}	{0.23}	{0.01}	{0.00}	{0.00}
F_{STAB} (Model 2)	{0.05}	{0.08}	{0.11}	{0.06}	{0.08}	{0.23}
$\chi^2_{LM}(14)$ (for Model 3 variables)	{0.00}	{0.00}	{0.00}	{0.81}	{0.05}	{0.16}
$\chi^2_{LM}(4)$ (Model 3: $(t p_t^f -_{t-1} p_{t-1})$ and $(t y_t^f -_{t-1} y_{t-1})$)	{0.01}	{0.00}	{0.00}	{0.47}	{0.13}	{0.53}
$\chi^2_{LM}(4)$ (Model 3: $(t p_{t+1}^f -_{t-1} p_t^f)$ and $(t y_{t+1}^f -_{t-1} y_t^f)$)	{0.00}	{0.08}	{0.02}	{0.85}	{0.86}	{0.99}
$\chi^2_{LM}(6)$ (for Model 3: sp_t)	{0.00}	{0.00}	{0.00}	{0.49}	{0.01}	{0.02}
F_{STAB} (Model 3)	{0.51}	{0.17}	{0.28}	{0.26}	{0.04}	{0.48}

Notes: Standard errors are given in (.). R^2 is the squared multiple correlation coefficient, $\hat{\sigma}$ the standard error of the regression and $F_{SC(4)}$, F_{FF} , F_H , F_N and F_{STAB} report p-values in {·} for F-test statistics for serial correlation (SC), functional form (FF), normality (N) and heteroscedasticity (H) and a Chow test of the stability of regression coefficients. The $\chi^2_{LM}(14)$ gives p-values in {·} from a chi-squared test statistic (with 14 d.f.) for the joint test of zero restrictions on the coefficients of two lags each of forecasts of inflation and output growth $(t p_t^f -_{t-1} p_{t-1})$, $(t y_t^f -_{t-1} y_{t-1})$, $(t p_{t+1}^f -_{t-1} p_t^f)$ and $(t y_{t+1}^f -_{t-1} y_t^f)$, provided by the SPF, and of six lags of the spread sp_t . The remaining χ^2_{LM} statistics provide a breakdown of the contribution of each of these respective variables in Model 3.

Table 3: RMSE's and Diebold-Mariano Statistics

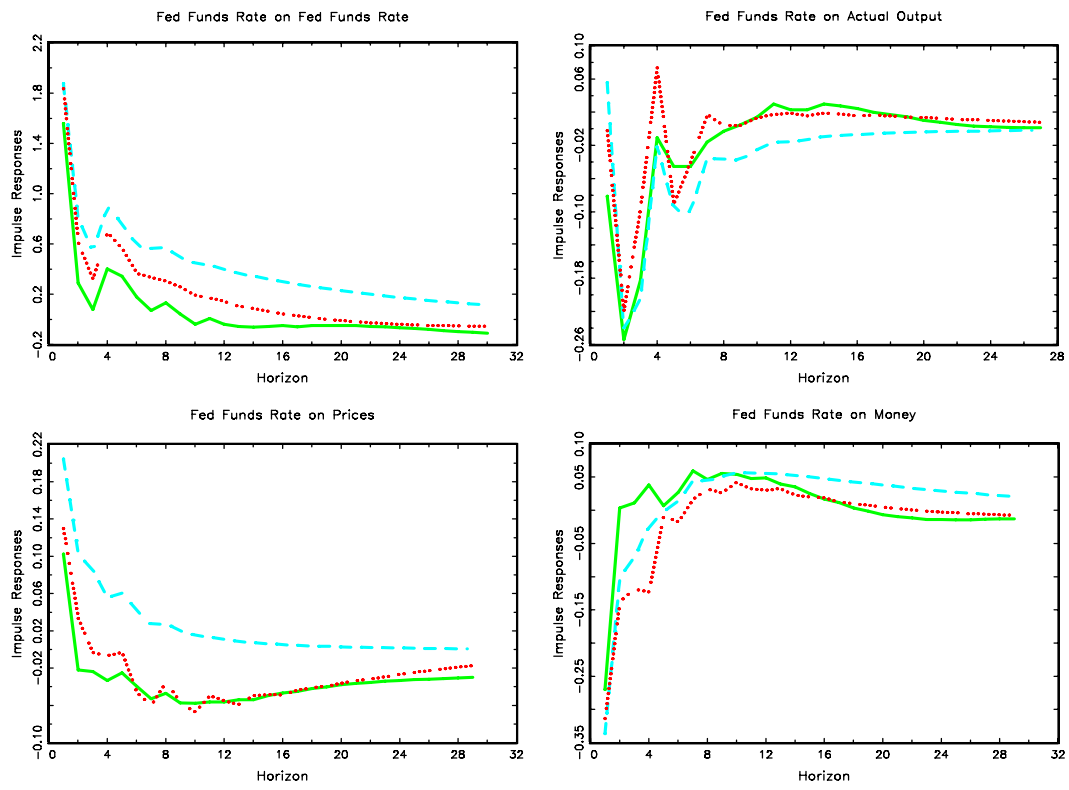
	RMSE's			Diebold-Mariano Statistics	
	Model 1'	Model 2	Model 3	Model 1' vs 2	Model 2 vs 3
1. $\sqrt{\frac{1}{N-1} \sum_{\tau=1}^{N-1} (\tau+1y_{\tau} - \widehat{\tau+1y_{\tau}})^2}$	0.0104	0.0073	0.0063 ^{††}	3.2374 [0.002]	2.3285 [0.022]
2. $\sqrt{\frac{1}{N-3} \sum_{\tau=1}^{N-3} (\tau+3y_{\tau} - \widehat{\tau+3y_{\tau}})^2}$	0.0122	0.0094	0.0091 ^{††}	2.3241 [0.022]	0.2881 [0.774]
3. $\sqrt{\frac{1}{N-5} \sum_{\tau=1}^{N-5} (\tau+5y_{\tau+2} - \widehat{\tau+5y_{\tau+2}})^2}$	0.0174	0.0182	0.0159 ^{††}	-0.5315 [0.596]	1.5055 [0.136]
4. $\sqrt{\frac{1}{N-7} \sum_{\tau=1}^{N-7} (\tau+7y_{\tau+4} - \widehat{\tau+7y_{\tau+4}})^2}$	0.0226	0.0261	0.0238 ^{††}	-1.9892 [0.050]	0.9712 [0.334]
5. $\sqrt{\frac{1}{N-1} \sum_{\tau=1}^{N-1} (\tau+1p_{\tau} - \widehat{\tau+1p_{\tau}})^2}$	0.0068	0.0069	0.0040 ^{††}	-0.7868 [0.434]	3.4783 [0.001]
6. $\sqrt{\frac{1}{N-1} \sum_{\tau=1}^{N-1} (\tau+4p_{\tau+3} - \widehat{\tau+4p_{\tau+3}})^2}$	0.0043	0.0044	0.0031 ^{††}	-1.1519 [0.253]	2.9863 [0.004]
7. $\sqrt{\frac{1}{N} \sum_{\tau=1}^N (x_{\tau}^r \Omega_{\tau} - \tilde{x}_{\tau}^f \Omega_T)^2}$		0.0082	0.0075 ^{††}		
8i. $\sqrt{\frac{1}{N} \sum_{\tau=1}^N (g_{\tau}^r \Omega_{\tau} - g_{\tau}^f \Omega_T)^2}$ for $\lambda = 0.1$		1.88×10^{-5}	1.74×10^{-5}		
8ii. $\sqrt{\frac{1}{N} \sum_{\tau=1}^N (g_{\tau}^r \Omega_{\tau} - g_{\tau}^f \Omega_T)^2}$ for $\lambda = 0.3$		0.0018	5.22×10^{-5} ^{††}		
8iii. $\sqrt{\frac{1}{N} \sum_{\tau=1}^N (g_{\tau}^r \Omega_{\tau} - g_{\tau}^f \Omega_T)^2}$ for $\lambda = 0.5$		0.0018	8.69×10^{-5} ^{††}		

Notes: The table reports RMSE and Diebold-Mariano statistics for the model specifications described in the text. $\tilde{x}_{\tau}^r | \Omega_{\tau}$ and $\tilde{x}_{\tau}^f | \Omega_T$ respectively denote the real-time and final output gap, as described in the text, and $g_{\tau} | \Omega_{\tau} = \lambda(\tilde{x}_{\tau} | \Omega_{\tau}) + (\tau+1p_{\tau} - \tau p_{\tau-1})^2$. The models are estimated for $t = 1968q4, \dots, \tau, \tau = 1985q4 - 2008q4$ and $T = 80$. The statistics in square brackets denote p-values. The [†] and ^{††} denote the results of the test that the difference between the RMSE of Model 2 and 3 are the same under the null that the data is generated under Model 2; the symbols denote significance at the 10% and 5% levels, respectively.

Table 4: Model 3 Nowcast and Forecast Conditional Event Probabilities of NBER-dated Contractions

	1990q3	1990q4	1991q1	1991q2	1991q3	2001q1	2001q2	2001q3	2001q4	2008q4
Nowcast ($h = 0$)										
$[sp_t], [({}_t p_t^f - {}_t p_{t-1}), ({}_t y_t^f - {}_t y_{t-1})], [({}_t p_{t+1}^f - {}_t p_t^f), ({}_t y_{t+1}^f - {}_t y_t^f)]$	0.5503	0.4774	0.9134	0.6887	0.0933	0.3520	0.4249	0.7392	0.9712	0.9971
sp_t	0.1446	0.1012	0.5404	0.3480	0.0954	0.2087	0.2736	0.3246	0.0797	0.2430
$({}_t p_t^f - {}_t p_{t-1})$ and $({}_t y_t^f - {}_t y_{t-1})$	0.5373	0.7871	0.9688	0.7097	0.1136	0.4695	0.3975	0.7915	0.9885	0.9984
$({}_t p_{t+1}^f - {}_t p_t^f)$ and $({}_t y_{t+1}^f - {}_t y_t^f)$	0.4261	0.3664	0.7758	0.4557	0.1269	0.2250	0.2210	0.3876	0.2528	0.8350
Four period ahead forecast ($h = 4$)										
$[sp_t], [({}_t p_t^f - {}_t p_{t-1}), ({}_t y_t^f - {}_t y_{t-1})], [({}_t p_{t+1}^f - {}_t p_t^f), ({}_t y_{t+1}^f - {}_t y_t^f)]$	0.6373	0.2379	0.2681	0.2379	0.1682	0.1481	0.1783	0.2152	0.1835	0.1093
sp_t	0.8238	0.2718	0.2359	0.2248	0.1525	0.1330	0.1554	0.2625	0.2099	0.0811
$({}_t p_t^f - {}_t p_{t-1})$ and $({}_t y_t^f - {}_t y_{t-1})$	0.1386	0.0969	0.1006	0.1813	0.0906	0.1090	0.0799	0.0664	0.0498	0.1047
$({}_t p_{t+1}^f - {}_t p_t^f)$ and $({}_t y_{t+1}^f - {}_t y_t^f)$	0.2604	0.2394	0.2246	0.3129	0.3131	0.2993	0.2271	0.2182	0.2824	0.3195

Notes: The table reports nowcast event probabilities of NBER-dated contractions, conditioning on information sets consisting of various combinations of the forward-looking variables. The variables are categorised into three sets of information set, namely, the spread, ${}_t sp_t$, the SPF time- t forecasts of inflation and output growth, $({}_t p_t^f - {}_t p_{t-1})$, $({}_t y_t^f - {}_t y_{t-1})$, and the time- $(t + 1)$ SPF forecasts of output growth and inflation, $({}_t p_{t+1}^f - {}_t p_t^f)$, $({}_t y_{t+1}^f - {}_t y_t^f)$. The information sets listed in the table implies their inclusion in the respective simulation experiment. The simulation experiment underlying the computation of these probabilities is the same as that for the nowcast probabilities plotted in Figure 3, and is as detailed in the text.



----- Model 1 - - - - - Model 2 - - - - - Model 3

Figure 1: Impulse Responses of a Federal Funds Rate Shock

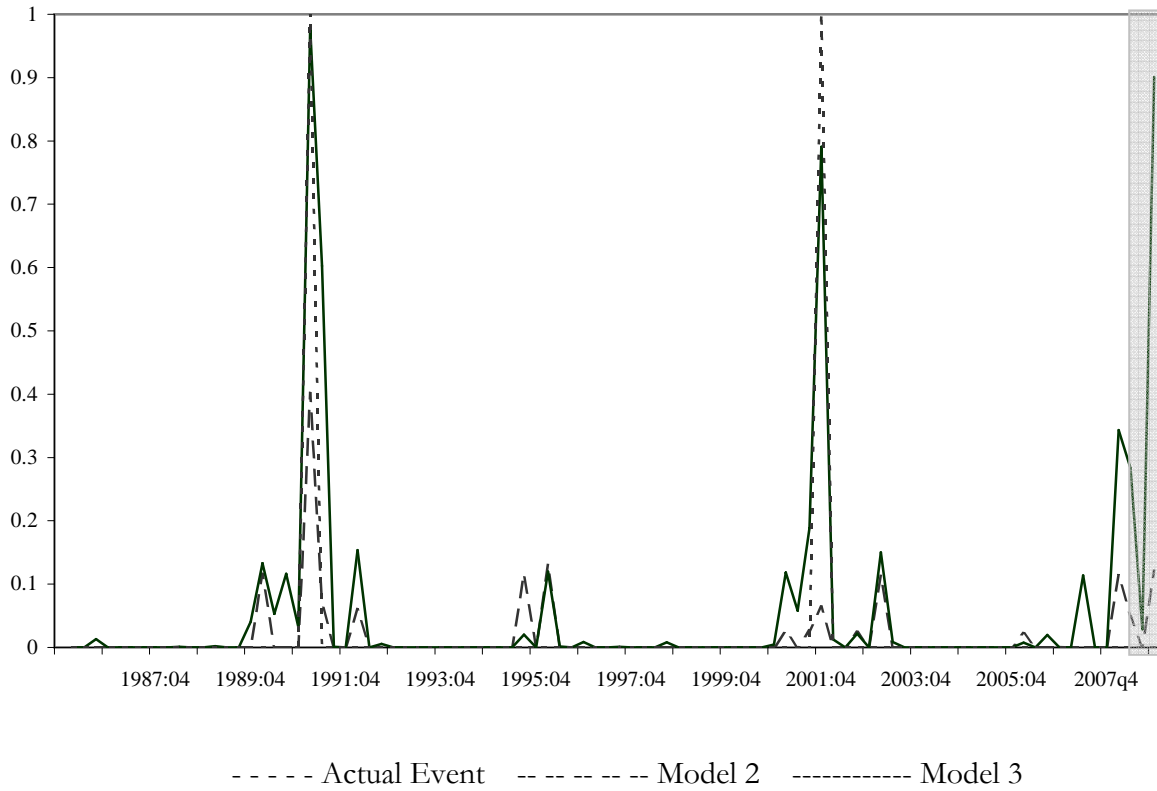


Figure 2: "Nowcast" probabilities of two periods of consecutive negative growth;

$$pr \{ [(t+2)y_{t-1} - t+1 y_{t-2}) < 0] \cap [(t+3)y_t - t+2 y_{t-1}) < 0] \}$$

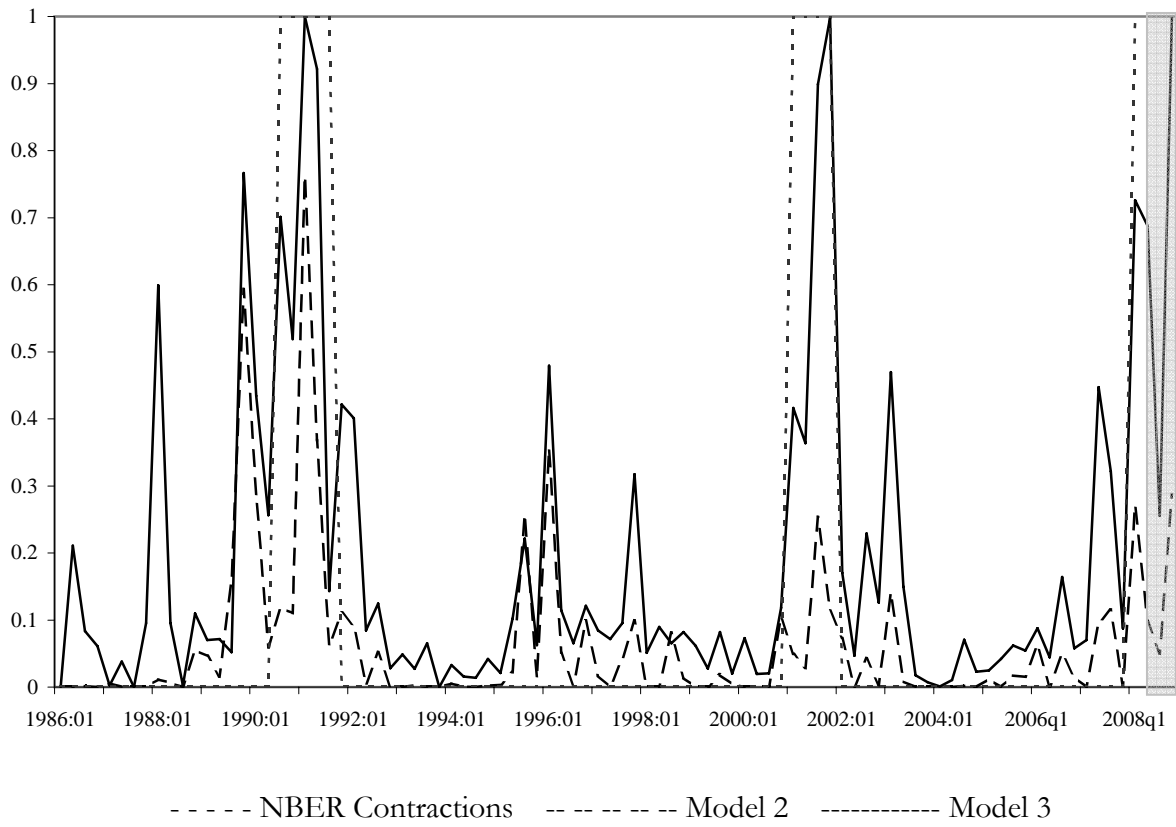


Figure 3: "Nowcast" probabilities of NBER Periods of Contraction

References

- Batchelor, R. (1986), “Quantitative versus Qualitative Measures of Inflation Expectations”, *Oxford Bulletin of Economics and Statistics*, Vol. 48, pp. 99-120.
- Bernanke, B.S and I. Mihov (1998), “Measuring Monetary Policy”, *Quarterly Journal of Economics*, 113, 869-902.
- Blinder, A., (1997), “What Central Bankers Could Learn from Academics - And Vice-Versa”, *Journal of Economic Perspectives*, 11, 2, pp. 3-19.
- Bordo, M. D and J. G. Haubrich (2008), “The Yield Curve as a Predictor of Growth: Long-Run Evidence, 1875–1997”, *The Review of Economics and Statistics*, 90, 182-185.
- Brunner, A.D. (2000), “On the Derivation of Monetary Policy Shocks: Should We Throw the VAR Out with the Bath Water?”, *Journal of Money Credit and Banking*, 32, 254-279.
- Christiano, L.J., M. Eichenbaum and C.L. Evans (1999), “Monetary Policy Shocks: What Have We Learned and to What End?” Chapter 2 in J.B Taylor and M. Woodford (eds.), *Handbook of Macroeconomics*, Volume 1A. North-Holland, Elsevier: Amsterdam.
- Clark, T. and M. McCracken (2001), “Tests of Equal Forecast Accuracy and Encompassing for Nested Models”, *Journal of Econometrics*, Vol. 105, pp. 85-110.
- Clements, M. and D. Hendry (2005), “Information in Economic Forecasting”, *Oxford Bulletin of Economics and Statistics*, 67 (Supplement), 713-753.
- Cogley, T. J. and T. Sargent (2005), “The conquest of US inflation: Learning and Robustness to Model Uncertainty”, *Review of Economic Dynamics*, Vol. 8(2), pp. 528-563.
- Croushore, D. (2008), “Frontiers of Real-Time Data Analysis”, *Working Paper*, University of Richmond, March.

- Croushore, D. and C. Evans (2006), “Data Revisions and the Identification of Monetary Policy Shocks”, *Journal of Monetary Economics*, Vol. 53, pp. 1135–1160.
- Diebold, F. X. and R. S. Mariano (1995), “Comparing Predictive Accuracy”, *Journal of Business and Economic Statistics*, Vol. 13, pp. 253-265.
- Estrella, A. and M. Trubin, (2006), “The Yield Curve as a Leading Indicator: Some Practical Issues”, *Current Issues in Economics and Finance*, Vol. 12, No. 5, pp. 1-7.
- Evans, G.W. and S. Honkapohja, (2001), *Learning and Expectations in Economics*, Princeton University Press, Princeton, NJ.
- Gali, J. and M. Gertler (2007), “Macroeconomic Modelling for Monetary Policy Evaluation”, *Journal of Economic Perspectives*, vol. 21, 4, pp. 25-46.
- Garratt, A., K.C. Lee, E. Mise and K Shields (2008), “Real Time Representations of the Output Gap”, *The Review of Economics and Statistics*, forthcoming.
- Garratt, A., K. C. Lee, M.H. Pesaran and Y. Shin (2003a), “A Long Run Structural Macroeconometric Model of the UK”, *Economic Journal*, 113, 487, 412-455.
- Garratt, A., K. C. Lee, M.H. Pesaran and Y. Shin (2003b), “Forecast Uncertainty in Macroeconometric Modelling: An Application to the UK Economy”, *Journal of the American Statistical Association*, 98, 464, 829-838.
- Garratt, A., K.C. Lee, M.H. Pesaran and Y. Shin, (2006), *Global and National Macroeconometric Modelling: A Long-Run Structural Approach*, Oxford University Press, Oxford.
- Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.
- Jacobs, J. P. A. M and S. van Norden (2006) “Modeling Data Revisions: Measurement Error and Dynamics of “True” Values”, CCSO Working Paper 2006-07.
- Kim, K. and A.R. Pagan (1995), “The Econometric Analysis of Calibrated Macroeconomic Models,” Chapter 7 in *Handbook of Applied Econometrics: Macroeconomics*. edited by M.H. Pesaran and M. Wickens. Oxford: Basil Blackwell.

- Koenig, E., S. Dolmas and J. Piger (2003), “The Use and Abuse of Real-Time Data in Economic Forecasting”, *Review of Economics and Statistics*, 85, 618-628.
- Koop, G., M.H. Pesaran and S.M. Potter (1996), “Impulse Response Analysis in Nonlinear Multivariate Models,” *Journal of Econometrics*, 74, 119-47.
- Lee, K.C. (1994), Formation of Price and Cost Inflation Expectations in British Manufacturing Industries: A Multisectoral Analysis, *Economic Journal*, 104, 372-386.
- Lee, K.C. and K. Shields (2000), “Expectations Formation and Business Cycle Fluctuations; An Empirical Analysis of Actual and Expected Output in UK Manufacturing, 1975-1996”, *Oxford Bulletin of Economics and Statistics*, 62, 4, 463-490.
- Mankiw, G. and M.D. Shapiro (1983), “News or Noise: An Analysis of GNP Revisions”, *Survey of Current Business*, 66, 20-25.
- Mankiw, G., R. Reis and J. Wolfers (2003), “Disagreement about Inflation Expectations”, *NBER Macroeconomic Annual*, 209 - 248.
- Orphanides, A. (2001), “Monetary Policy Rules Based on Real-Time Data”, *American Economic Review*, 91, 964-985.
- Orphanides, A., R.D. Porter, D. Refschneider, R. Tetlow and F. Finan (2000), “Errors in the Measurement of the Output Gap and the Design of Monetary Policy”, *Journal of Economics and Business*, 52, 117-141.
- Orphanides, A. and J. C. Williams (2002), Imperfect Knowledge, Inflation Expectations and Monetary Policy, NBER Working Paper 9884.
- Pesaran, M. H. and R. Smith (2006), “Macroeconomic Modelling with a Global Perspective”, *The Manchester School*, Vol. 74, Supplement 1, pp. 24-49.
- Roberts, J. (1995), “New Keynesian Economics and the Phillips curve”, *Journal of Money, Credit and Banking*, 27, 975-984.

Roberts, J. (1997), "Is Inflation Sticky?", *Journal of Monetary Economics*, 39, 2, 173-196.

Rotemberg, J. and M. Woodford (1999), "Interest Rate Rules in an Estimated Sticky Price Model", in Taylor J.B. (ed.) *Monetary Policy Rules*, 57-119. University of Chicago Press: Chicago.

Sims, C. and T. Zha (1998), "Bayesian Methods for Dynamic Multivariate Models", *International Economic Review*, 39, 949-968.

Smith, J. and McAleer, M. (1995), "Alternative Procedures for Converting Qualitative Response Data to Quantitative Expectations: An Application to Australian Manufacturing", *Journal of Applied Econometrics*, 10, 2, 165-185.