"Keep It Real!": A Real-time UK Macro Data Set

Don M. Egginton
Daiwa Institute of Research Europe

Andreas Pick
University of Cambridge

Shaun P. Vahey*
University of Cambridge

First draft: 30 June 2001
This draft: 3 September 2001

Abstract

We present a real-time macro data set for the UK. Each variable has many different vintages—reflecting the revisions and updates that occur over time. Our aim is to provide a resource for researchers evaluating UK forecasting performance and policy-making in real time. We illustrate the importance of these data by analysing their impacts on UK inflation forecasts and monetary policy in the late 1980s. We find that, contrary to the view of contemporary policy-makers, the initial measurements of demand-side macro variables did not disguise inflationary pressures.

Keywords: Data revisions; Real-time data; Inflation forecasts

JEL classification: C8; E37; E52

*Correspondence to: Shaun Vahey, Faculty of Economics and Politics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge, UK, CB3 9DD. E-mail: shaun.vahey@econ.cam.ac.uk, Fax: +44 (0)1223 335475. We thank David Hendry, Simon Price, Ken Wallis, Rod Whittaker and participants at the Warwick Macro Modelling Seminar 2001 for helpful comments. We are grateful to Mark Green, Elena Loukoianova and Patricio Feres for research assistance. The real-time macroeconomic data are available from http://www.econ.cam.ac.uk/dae/keepitreal/. Partial financial support from the ESRC (Research Grant No. L38251021) is acknowledged gratefully.

1 Introduction

Nearly all UK macro research is based on the "most recent vintage" of data. That is, for a given variable, the researcher uses the most recently measured values. As the official data producing agency—the Office for National Statistics (ONS)—revises the figures, successive researchers, each using the most recent vintage, analyse different data. Furthermore, the data used by researchers for forecast and policy evaluation differ from those confronting forecasters and policy-makers in real time.

The literature examining the impact of data revisions has a long tradition including Zellner (1958), Stekler (1967) and Harvey et al. (1983). A series of papers by Patterson and Hervai (1991b, 1991a) and Patterson (1994, 1995) have examined the relationship between different vintages of UK data for a variety of time series. Real-time data have been used by a number of researchers for policy analysis including Orphanides and van Norden (1999) and Orphanides et al. (2000) (output gaps), and Orphanides (2000) (Taylor rules).

The only existing source of multiple variable real-time data that we are aware of refer to the US and were compiled by the Federal Reserve Bank of Philadelphia. The data are available at http://www.phil.frb.org/econ/forecast/; data descriptions are provided by Croushore and Stark (1999, 2000). In this paper, we present a complementary real-time macro data set for the UK. The data can be downloaded from http://www.econ.cam.ac.uk/dae/keepitreal/, together with detailed descriptions. We demonstrate the importance of using real-time data by examining the impacts on UK inflation forecasts and monetary policy in the late 1980s.

2 Data details

Typically, macro data sets contain only the most recent vintage of a given time series variable, with the data recorded as a column (or row) vector. In our real-time data set, for each variable, the data are stored as a matrix. Each successive column vector of the matrix represents a new vintage of data, containing the most recent measurements available at that vintage date. The data are stored by variable and were collected by examining various issues of *Economic Trends* and *Financial Statistics*, both of which are published by the ONS (formally the Central Statistical Office). The figures reported were in the public domain at the end of the month in question. For each vintage, the observations are identical to those in the relevant published source.[1]

The variables comprise: (real) GDP, (real) consumers' expenditure, unemployment, (real) industrial production, retail sales (volume), monetary aggregates (M0, M3, M4), budget deficit (public sector borrowing requirement) and average earnings. The first two variables are quarterly; the remainder are monthly. With the exception of the money figures, the first vintage is January 1980 and the last is

[1] Unfortunately, the window length reported by the source publications is affected by page layout considerations—it varies by variable and by vintage date. Missing data are recorded as empty cells.

June 1999. Reflecting availability in the source publications, for M0, M3 and M4 the first vintages are June 1981, January 1980 and June 1987 and the last vintages are June 1999, August 1989 and June 1999 respectively. All (vintage dependent) ONS codes are available from the website.

The size and number of revisions to UK macro data have been sources of much concern, particularly during the late 1980s. An official scrutiny into the quality of data was conducted in 1988. The "Pickford Review" proposed a number of reforms designed to improve the quality of macro statistics, many of which were adopted over the next decade (see Wroe (1993) for details). Kenny and Rizki (1992) and Wroe (1993) document the extent of revisions to various ONS series. A number of researchers, including Johnson (1991) and Smith (1992) have argued that the quality of UK macro statistics deteriorated in the late 1980s, with initial measurements of key variables subject to downward bias. Policy-makers at both HM Treasury and the Bank of England felt that this bias—something more than previously experienced mean zero measurement error—delayed the tightening of monetary policy in the late 1980s, contributing to the inflationary boom (see, for example, Hibberd (1990) and Leigh-Pemberton (1990)).

3 A case study: UK inflation forecasts in the late 1980s

In order to assess the impact of using real-time data on inflation forecasts, we utilise two distinct data sets, distinguished by the vintages used but containing common macro indicators. The first is based on the most recent vintage available currently (July 2001) for each variable. Henceforth, we refer to these data as "current vintage". The second is based on the most recent vintages available in real time—we refer to these data as "real-time vintage".

We restrict our consideration to the following base set of macro indicators (where Δ denotes the year-on-year growth rate)[2]: INFL (inflation), ΔFT (FT all-share index), ΔEM (£-DM exchange rate), ΔED (£-US$ exchange rate), USSP (US yield spread), ΔTBILL (change in UK 3-month treasury bill rate), ΔHP (house prices), CP (commodity prices), ΔPPI (producer output prices), ΔU (unemployment), ΔIP (real industrial production), ΔM0 (money supply), ΔGDP (real gross domestic product), and ΔC (real consumers' expenditure). Full details of the variables used, including the ONS mnemonics, can be obtained from the website, together with the data themselves. All variables are available monthly except consumers' expenditure and GDP which are available quarterly. To obtain monthly series for these variables we interpolate linearly (but not extrapolate) the quarterly year-on-year growth rates using the nearest quarterly observations.

We generate out-of-sample inflation forecasts for the late 1980s by utilising a recursive modelling strategy (Pesaran and Timmermann 1995). Our approach is based on the view that inflation can be forecast by exploiting statistical relationships with a variety of macro indicators and that these relationships may vary over

[2]Clearly, the set of variables that can be used to forecast inflation is very large. For example, Wallis et al. (1987) describes the (quarterly) HM Treasury Model as having 700 equations and 1275 variables. For computational reasons, we use a (much) more parsimonious base set.

the business cycle. Given the well-documented economic reforms and switches in intermediate macro targets in the period (see, for example, Johnson (1991)), it is very unlikely that one inflation forecasting model held throughout.[3]

We model a typical forecaster searching for a satisfactory empirical specification among the many believed to be useful for predicting inflation over the period from April 1982 to October 1990. We use observations from April 1982 to December 1986 as a "training period"; and make 12-month ahead forecasts for the period January 1988 to October 1990. The start of the sample was chosen to exclude the early 1980s recession; the end reflects the resignation of Nigel Lawson as Chancellor and an associated change in monetary policy in October 1989.

Using the real-time vintage data, in every period, denoted τ, from January 1987 to October 1989, we estimate the following linear models by OLS:

$$\pi_t = \alpha_i + \beta_i' x_{i,t-12} + \varepsilon_{i,t}, \qquad i = 1, \ldots, 2^k, \quad t = 1, \ldots, \tau - 1,$$

where π_t is the 12-month growth rate of RPIX from $t-12$ to t.[4] The subscript i refers to the 16384 different models estimated in each period τ—one for each of 2^k combinations of the $k = 14$ macro indicators in the base set. Each model is identified by a unique combination of macro indicators included in $x_{i,t-12}$. All macro variables are contained within the information set at t–12, i. e. each variable is dated such that the most recent publicly available observation at $t - 12$ is included. In every period, τ, we use the Akaike Information Criterion (AIC) to select a preferred model from the many models estimated.[5] The resulting optimal models, one for each period from January 1987 to October 1989, are used to construct our forecasts for January 1988 to October 1990 ($\pi_{\tau+12}$).

For comparison, we repeat this recursive model selection and forecast procedure using current (rather than real-time) vintage data. The resulting two 12-month ahead forecast series, one for each data set, are shown in Figure 1. For most of the forecast period, the real-time forecasts exceed those based on the current vintage. That is, a forecaster using real-time data would have predicted stronger inflationary pressures than a researcher looking back at the period using the current vintage. At first pass this finding seems surprising. Kenny and Rizki (1992) and Wroe (1993) show that the initial measurements of a number of demand-side macro variables were biased downwards. Assuming that the estimated model parameters were the same for both data types, real-time vintage data would result in weaker evidence of inflationary pressure. In fact, the estimated parameters (and models) using the real-time vintage are very different from those using the current vintage. Notice also that by visual inspection, the real-time forecasts ex-

[3] A single more structural model, such as that in Hendry (2001), may produce superior forecasts; but, given that our focus is out-of-sample performance in the presence of structual change, it may not.

[4] The growth in RPIX is currently the Bank of England's preferred measure of underlying inflation. It excludes the component of the RPI involving mortgage interest payments. The impacts of revisions on RPI rather than RPIX inflation forecasts are similar. Details can be obtained from the authors on request.

[5] Results using other information criteria are similar and are available from the authors on request.

hibit greater volatility than the current vintage forecasts, reflecting the volatility in both the data and the estimated coefficients.

It is clear from Figure 1 that UK inflation was difficult to forecast out of sample in the late 1980s. Both forecasts series suggest (varying degrees of) forecast failure. It is interesting to note that a variety of independent forecasts and the public HM Treasury forecasts reported in Treasury Select Committee (1991) were similarly afflicted.[6]

Undoubtedly, inflation forecast failure contributed to the well-documented failure of monetary control in the late 1980s. Nigel Lawson, Chancellor in our forecast period, famously suggested in his 1985 Mansion House speech that "(t)he acid test of monetary policy is its record in reducing inflation ... The inflation rate is judge and jury" (Lawson 1992, p.481). Yet RPIX inflation—Lawson's preferred measure of underlying inflation (Lawson 1992, p.480)—increased steadily from July 1986 until Lawson's resignation in October 1989. Figure 1 suggests that contrary to the contention of Hibberd (1990) and Leigh-Pemberton (1990), the downward bias in the initial measurements of demand-side variables did not disguise inflationary pressures in the late 1980s. Quite the reverse; using real-time data, for most of the period under consideration, the evidence of lack of monetary discipline was somewhat stronger.

This episode in UK monetary policy can be used to illustrate the importance of using real-time data for policy evaluation. Consider a researcher assuming that Lawson followed an RPIX inflation forecast target of 4.5 percent or less (the rate prevailing when he started as Chancellor in July 1983). If our researcher believed that Lawson could reasonably have made the forecasts shown in Figure 1, then she would conclude that policy errors were made whenever the 12-month ahead forecasts exceeded the target. Using the current vintage forecasts, this occurred twice in our forecast period: between May 1987 and September 1987 (12-month ahead forecasts reported dated May 1988 and September 1988); and again from February 1989 to the end of the sample. Using the real-time vintage forecasts, the policy errors occurred later in the first case, July 1987 to October 1987 (arguably, Febuary 1988), and earlier in the second case, October 1988 to the end of the sample.[7]

Of course, in reality, non-inflation related objectives (at times) played a part in Lawson's setting of monetary policy. The candidates include growth (in particular, avoiding a recession following the 1987 stock market crash), tax reforms and exchange rates targets. Since our aim is to demonstrate that real-time data matter for forecast and policy evaluations, we leave these issues to be explored in a more complete real-time study of UK monetary policy in the late 1980s.

[6]The forecast failure of our models is driven partly by the assumption that inflation is stationary in the (short) period under consideration. Ex post, it is clear that modelling inflation as non-stationary improves forecasting performance but, arguably, such an assumption is inconsistent with the notion of real-time monetary control.

[7]Actual base rates were raised repeatedly in the second half of 1988 from 8.9 percent to 12.9 percent. This continued though 1989 to a peak of 14.9 percent. Apparently this was insufficient to offset inflationary pressures.

References

Croushore, D., and T. Stark (1999) 'A real-time data set for macroeconomists: Does the data vintage matter?' Federal Reserve Bank of Philadelphia *Working Paper 99–21.*

— (2000) 'A real-time data set for macroeconomists: Does the data vintage matter for forecasting?' Federal Reserve Bank of Philadelphia *Working Paper 00–6.*

Harvey, A.C., C.R. McKenzie, D.P.C. Blake, and M.J. Desai (1983) 'Irregular data revisions.' In *Applied Time Series Analysis of Economic Data,* ed. A. Zellner (Washington DC) pp. 329–39.

Hendry, D.F. (2001) 'Modelling UK inflation, 1875–1991.' *Journal of Applied Econometrics* 16, 255–276.

Hibberd, J. (1990) 'Official statistics in the late 1980s.' *Treasury Bulletin* Summer.

Johnson, C. (1991) *The Economy Under Mrs Thatcher 1979–1990* (London: Penguin).

Kenny, P.B., and U.M. Rizki (1992) 'Testing for bias in initial estimates of economic indicators.' *Economic Trends* 463, 77–86.

Lawson, N. (1992) *The View from No. 11* (London: Bantam Press).

Leigh-Pemberton, R. (1990) 'Monetary policy in the second half of the 1980s.' *Bank of England Quarterly Bulletin* pp. 215–220.

Orphanides, A. (2000) 'An activist stabilization policy and inflation: The Taylor rule in the 1970s.' mimeo, Board of Governors of the Federal Reserve System.

Orphanides, A., and S. van Norden (1999) 'The reliability of output gap estimates in real time.' mimeo, Board of Governors of the Federal Reserve System.

Orphanides, A., R.D. Porter, D. Reifschneider, 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, 117–41.

Patterson, K.D. (1994) 'A state space model for reducing the uncertainty associated with preliminary vintage of data with an application to aggregate consumption.' *Economics Letters* 46, 215–222.

— (1995) 'A state space approach to forecasting the final vintage of revised data with an application to the index of industrial production.' *Journal of Forecasting* 14, 337–350.

Patterson, K.D., and S.M. Hervai (1991a) 'Are different vintages of data on the components of GDP co-integrated? Some evidence for the United Kingdom.' *Economics Letters* 35, 409–413.

— (1991b) 'Data revisions and the expenditure components of GDP.' *Economic Journal* 101, 997–901.

Pesaran, M.H., and A. Timmermann (1995) 'Predictability of stock returns: Robustness and economic significance.' *Journal of Finance* 50, 1201–1228.

Smith, D. (1992) *From Boom to Bust: Trial and Error in British Economic Policy* (London: Penguin).

Stekler, H.O. (1967) 'Data revisions and economic forecasting.' *Journal of the American Statistical Association* 62, 470–83.

Treasury Select Committee (1991) *Memoranda on Official Economic Forecasting* (London: HMSO).

Wallis, K.F., P.G. Fisher, J.A. Longbottom, D.S. Turner, and J.D. Whitley (1987) *Models of the UK Economy: A Fourth Review by the ESRC Macroeconomic Modelling Bureau* (Oxford: Oxford University Press).

Wroe, D. (1993) 'Improving macro-economic statistics.' *Economic Trends* 471, 191–199.

Zellner, Arnold (1958) 'A statistical analysis of provisional estimates of gross national product and its components, of selected national income components, and of personal saving.' *Journal of the American Statistical Association* 53, 54–65.

Figure 1: Inflation and forecasts