Discussion Paper No.19

Monetary Policy and Data Uncertainty: A Case Study of Distribution, Hotels and Catering Growth

by Lavan Mahadeva

External MPC Unit Discussion Paper No. 19*

Monetary Policy and Data Uncertainty: A Case Study of Distribution, Hotels and Catering Growth

By Lavan Mahadeva**

December 2007

ISSN: 1748 – 6203

I wish to thank Charlotta Groth, Chris Jeffery, Juan Carlos Parra and Andrew Sentance for their helpful comments. Any errors are mine.

(*) Disclaimer: These Discussion Papers report on research carried out by, or under supervision of the External Members of the Monetary Policy Committee and their dedicated economic staff. Papers are made available as soon as practicable in order to share research and stimulate further discussion of key policy issues. However, the views expressed are those of the authors and do not represent the views of the Bank of England or necessarily the views of other External Members of the Monetary Policy Committee.
(**) At the time of writing, External Monetary Policy Committee Unit, Bank of England, Threadneedle Street, London, EC2R 8AH. Currently working at the Banco de la República de Colombia.

Abstract

This paper is a case study of the real world monetary policy data uncertainty problem. The initial and the latest release for growth rates of the distribution, hotels and catering sector are combined with official data on household income and two surveys in a state-space model. Though important to the UK economy, the distribution, hotels and catering sector is apparently difficult to measure. One finding is that the initial release data is not important in predicting the latest release. It could be that the statistical office develop the initial release as a building block towards the final release rather than an estimate of it. Indeed, there is multicollinearity between the initial release and the retail sales survey, which would then contain the same early available information. A second finding is that the estimate of the later release is sensitive to the estimate of the average historical growth rate. This means that establishing priors for this parameter and testing for shift structural breaks should be very important.

Key words: Data Uncertainty, Distribution Sector, Kalman Filter, Monetary Policy.

1 Introduction

"Strange to know nothing, never to be sure of what is true or right or real."
Ignorance by Phillip Larkin, from The Whitsun Weddings

Monetary policymakers are not always sure what is true or right or real. The official data on important economic variables, such as consumption and output, can never be both accurate and timely. Initial releases are available up to quite recent quarters. But these get revised substantially and much later on, as it naturally takes time for the statistical office to bring in new information and better method.

Yet monetary policymakers still need to form some view of where important variables are. When they make forecasts, they need to determine some initial conditions. When they set interest rates, they need to know what they are reacting to.

The way in which they form these judgements does reflect that the official data is uncertain, though. Typically they take in information from more timely non-official sources, such as private sector surveys, and combine that with official sources to form a view. And they treat the view they form as a statistical estimate conditional on available information and assumption. It is acknowledged that the distribution as well as the mean of the estimate matters. And when new information and relevant information comes to light, the view can be revised.

A decision-making model with these features is straightforward to formalise. There is a well-known algorithm — the Kalman filter — that suits the purpose[1]. The Kalman filter allows us to combine different types of observed data with a theory that identifies the unobservable true series of interest. Combining an estimation objective; a definition and set of indicators; a set of initial conditions; and a numerical procedure, the filter produces an estimate of the true series within sample and a forecast of out-of-sample, with associated standard errors. The advantage of the filter is optimal under a wide class of problems of this type (see Anderson and Moore, 1994, for an explanation of how general the Kalman filter is). conform

In this paper I apply this method to a case study of estimating the UK distribution, hotels and catering sector´s output for the purpose of monetary policy-decision making. And there are some important differences when the data uncertainty exercise is applied as a tool for monetary policy.

Most importantly, monetary policy decision-making needs to be made transparent. Here I take that to mean that the outsider should be able not just to observe the numerical outcome of the data

[1] See for example, the exposition on the treatment of uncertainty in official data in Andrew Harvey's textbook (Harvey, 1989) on the Kalman filter.

uncertainty exercise, as say the estimate of the mean of the true variable of interest, but also to understand where that estimate came from. As I shall demonstrate, these estimates are likely to be quite sensitive to modelling assumptions. Under these circumstances, publishing a central estimate by itself does not convey much useful information. A transparent data uncertainty method would have to expose this conditionality.

Another set of peculiarities of the monetary policy data uncertainty problem come from the variety of data sources used:

- First there is the initial release of the official data on that series. For many national accounts output series, these initial estimates are available with a ten-week lag.
- Second there are many historical vintages of that same series. Later vintages are likely to be closer to the truth as more information and better method is bought to bear. But they can come too late to be of much use in the informing current conjuncture.
- A third category is other official data which might relate to the series of interest. For example official data on productivity might help predict wages (when combined with initial data on earnings). Account would have to be taken as to whether these data itself would be revised.
- A fourth source is data from other unofficial sources. This could be information from private-sector surveys. Surveys can be very timely. But typically they are not processed to the same rigour as late releases of official data. Another source is financial market data. The problem there is that they may not match exactly with the concept of interest. For example the financial market price of shares in sector can be available instantaneously, but can of course diverge from the current value-added output of the sector: the two are economically distinct concepts.
- A final, fifth, source of information is a prior understanding of how that economic variable is developing. A simple prior could be the historical mean growth rate of that series. More sophisticated priors could reflect theoretical knowledge of how the wider economy works, that is often hard to associate with a single data series.

A data set of this heterogeneity will typically be unbalanced. While survey data tend to have short histories, mature releases of official data are not available for recent quarters. The data set cannot just be topped and tailed to a common sample: monetary policymakers would need to combine data from these unbalanced series and not throw away potentially valuable information. The data set could also be heterogeneous in its frequency. Financial data is continuous, survey and indicator data are monthly, official data may be quarterly and some information may only be available once a year. Some data might only come irregularly. There can also be holes or missing observations where reporting temporarily broke down. In principle the method should be adjusted to cope with this.

Another problem arises because these data series all seek to measure very similar concepts. For that reason they may contain only the same common relevant information: they may be highly collinear. Multicollinearity need not affect the unbiasedness and minimum variance property of the Kalman filter estimate. But in small samples, it can make the estimates less reliable, in the sense that the estimated coefficients are highly sensitive to new observations, even in a large sample of data. The symptoms are instability in estimates over different samples, and large but offsetting parameters between two collinear indicators.

Then there are econometric problems that matter for any time-series estimation. There could be structural breaks. Changes are always afoot that are continually reshaping economic structures. Even if the economic structure does not change, the measurement technology may have altered. Of course, breaks can happen to the slope coefficients that multiply economic variables, but often a greater threat are breaks to the constants,[2] which can mean that underlying long-run growth rate of the series, or the statistical process, has shifted.

UK time series data exhibit heteroscedasticity, as the 1980s and early 1990s were typically more volatile than the decade that immediately followed. Heteroscedasticity can also arise in the data measurement, if measurement techniques become more precise over time. In principle the Kalman filter can provide for a very general type of heteroscedasticity: it can allow for error variances to time vary. But it might be better to put more structure on the heteroscedasticity, and allow it to develop in a systematic form.

And then there may be serial correlation to deal with. Again, there are two types of serial correlation: that to do with the underlying economic process and that to do with the measurement process. All the data has been seasonally adjusted but being from different sources, the series are subject to different methods of seasonal adjustment. This might imply some residual serial correlation in itself.

Ideally, then, the data uncertainty technology needs to be adapted to make estimates intuitive and transparent when it is used for monetary policy. It needs to take account of an unbalanced data set, to deal with possible structural breaks, to deal with serial correlation and heteroscedasticity, and to be well identified.

The paper presents a model of how monetary policy can make and explain decisions when data is uncertain. I have not been able to address all of the potential aspects of the monetary policy data uncertainty problem that I highlighted above. But I have dealt with some of the most important problems, and where not, I have at least exposed the conditionality of the estimates.

[2] See Clements and Hendry, 1999.

This paper is a case study. I apply the method to estimating value-added output growth in the distribution, hotels and catering sector as would be measured in the latest official release at least five quarters after the first release. The different aspects of my choice of case study needs some justification.

First why did I choose to tackle the distribution, hotels and catering sector, and not, for example, the whole of real GDP? One good reason is that most survey indicators apply to specific sectors, and at least in the UK, the economic behaviour across sectors seem diverse (see Oulton and Srinivasan, 2004, for example). Hence I would expect that more information would be gained about the true growth rate in GDP from combining separate estimates of sectoral output rather than from apply a data uncertainty model to the aggregate. Then distribution, hotels and catering is by no means small. Its share of GDP (value-added) was estimated to be 15.3% in 2003 (Office for National Statistics, 2006a). This is now as large as manufacturing sector (at 14.7%), on which many academic papers have been written. And, as I shall show below, it is not a well measured sector, compared to, for example, manufacturing.

Second, why did I choose to try and estimate the rate of growth of output and not a more theory-based concept, such as the component of output movements due to some economic shock? In part, this was because I wanted to keep the exercise more transparent. Consider if instead I had followed the objective of trying to estimate the demand shocks that drive both output and inflation. This theory-based alternative certainly gets closer to what matters for the monetary policy decision. After all, positive demand shocks generally imply both higher inflation and excessive bursts in output, and therefore monetary policy remit is to identify and act against them. And it would certainly be feasible to implement. Different vintages of observed data on output and inflation could be combined with a theory of potential output and a Phillips curve to estimate an output gap. One could adapt the single data vintage model of Kuttner (1994) to a real-time setting. This model would combine the data uncertainty problem with the problem of identifying the shocks and the transmission of monetary policy in one stage.

But I would argue that the outcome of estimating a theory-based objective is less transparent than trying to simply estimate output growth in some later release, simply because it would be more difficult to verify. One can only produce a series for demand shocks after much manipulation of the data, and by bringing in many assumptions about economic structure. Intricate subjectivity would creep into the calculation, and it would be more difficult to bring out the conditionality surrounding the estimate of the demand shock.

The simpler alternative is to seek to estimate only the later release of the growth rate. Of course, data on the later release of the growth rate of distribution hotels and catering sector output is not available in real time just as it isn't in the case of the demand shock. But one would only have to wait a year, and any releases of official data on that output measure in that intervening year should be moving closer to it. In this sense targeting later releases of the growth rate can be verified, and is therefore a more transparent objective, and is perhaps more suitable for communicating monetary policy. Of course this information would then have to be incorporated into a structural model in order to derive the monetary policy implications, as estimates of the initial conditions for where output is in the period where the forecast begins. But my presumption is that keeping the two stages separate is more transparent than trying to combine the data uncertainty exercise with the monetary policy problem.

And, third, why did I choose to cut off the latest data at five quarters, and not later? The data might get revised significantly even after five quarters. One could instead allow for the data to settle on a long-run final release over a horizon that is estimated rather than imposed, and use many different vintages to get there. Cunningham, Jeffery, Kapetanios and Labhard (2006) apply this approach to aggregate GDP growth, whose revisions settle very slowly. But later on I provide evidence that distribution hotels and catering sector output revisions are reasonably close (about two thirds of the way) to their long run after five quarters. A five-quarters cut-off guarantees that the final data has been through at least one Blue Book, which is the vintage by which the UK Office of National Statistics rebalance the accounts and undertake a first substantial revision. And finally, aiming to predict very late releases, for example after ten years, would mean that an outsider would have wait long to verify the estimate. So the choice was also made again with the cost of transparency in mind.

A related choice is over how far beyond the data set we want to forecast and backcast. I would argue that the relevance of this type of model for monetary policy will typically be limited to a window at most beginning up to a year or two ago and no further than the year ahead. Given that the model is directed at estimating the true growth rate of output, and is therefore not a structural model, anything that it can tell us about what happened in the distant past is not likely to matter much for monetary policy now. Similarly it is doubtful that a model of this type will add more useful information to a monetary policy forecast beyond a year ahead. The window on which I have focussed begins five quarters ago and ends a year ahead.

In the next section, Section 2, I explain the different sources of information that might be relevant to estimating true distribution, hotels and catering output, including the definition of the true variable of interest. Section 3 lays out the state space model that is used to estimate the distribution of DHC growth and explains how that was estimated. Section 4 explains the estimation procedure and Section 5 discusses the parameter estimates. Section 6 looks at the diagnostics of the model. Section 7 reveals the estimated DHC growth series. There I break down the estimated mean into the contributions of the data series. In Section 8 I decompose in terms of variances. In Section 9 I test for multicollinearity, and in Section 10 for robustness due to estimation uncertainty. Section 11 concludes.

2 Distribution Sector: A Case Study

2.1 The Official Data

To set the scene, Chart 1 shows the vintage of official data on the value-added output of the distribution, hotel and catering sector (henceforth DHC) that was available in early 2006. The pattern is of a sharp slowdown in 2003 and 2004 followed by a recent recovery. But that the recovery is recent means that it is present in only initial official releases for the series. A question that one could ask is whether that recovery might eventually be revised away.

We would worry about this because we know that initial releases of DHC output tend to be poor predictors of later, final releases. Chart 2 compares the two. In defining the final release I took the latest vintage and cut off any data that has not been through five quarters of revisions, as justified in the introduction. Chart 2 shows that although there is some broad relationship between the two, it is by no means exact.

Chart 1: DHC output: latest release

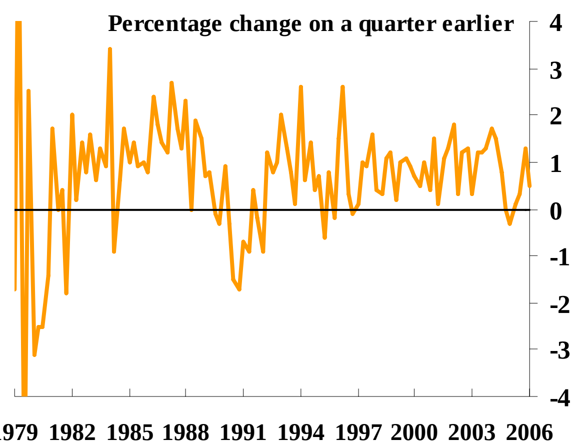
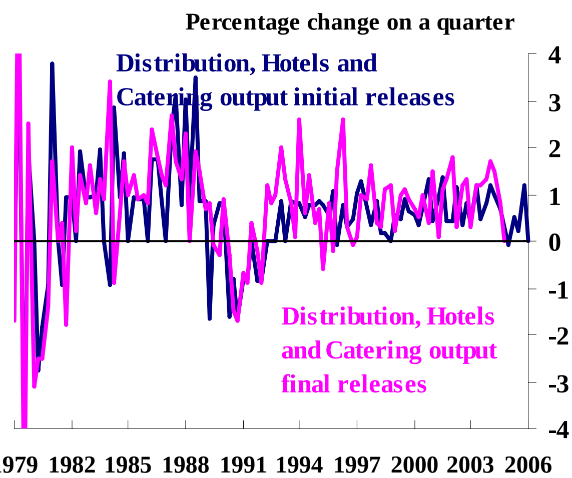


Chart 2: DHC output: first and final releases



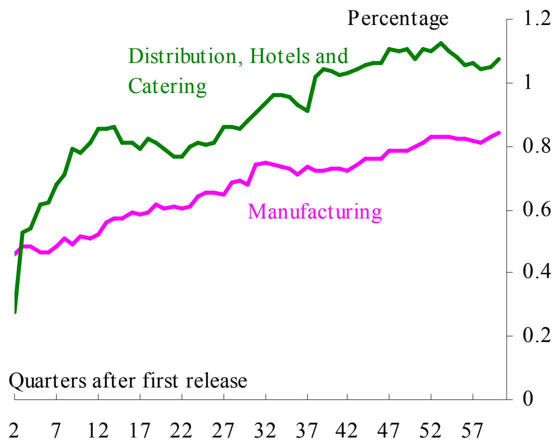
To explore this more formally, I regressed the initial release on the final release. The form of the regression is fairly simple: pre-testing with OLS estimates revealed that there was no role for a lagged final releases but a lagged initial release term was needed. I carried out a GARCH(1) estimation as pre-testing also identified some need to allow for some heteroscedasticity in the residuals. Diagnostic testing of the residuals indicated that the estimation was well specified: an ARCH test indicated no further heteroscedasticity, and a Q statistic of the partial autocorrelations of the residuals indicated no serial correlation of either short or long lags.

Table 1: Regression of final official data on the first release and a constant		
	Mean coefficient estimate	Standard deviation
Constant	0.31	0.12**
Initial release	0.49	0.09**
Initial release (-1)	0.18	0.09**
Garch regression		
Constant	0.07	0.03**
Garch (-1)	0.87	0.04**
$R^2 = 0.35$; $\overline{R}^2 = 0.32$; standard error = 1.17%; Sample 1979Q2-2004Q4 (103 observations)		

Table 1 reports the results. If the first release were a very good predictor of the final release we could expect the long run slope coefficient of that relationship to be one; the long-run constant to be zero; and the residuals to have a small variance. This is not what the estimates indicate. The sum of the coefficients on the current and lagged initial release terms is significantly less than one (a Wald test soundly rejected this null); there is a significant bias of average 0.3pp; and the estimated equation standard errors are large (even larger than the mean of final release DHC growth). The evidence is then that the initial estimate is an imperfect predictor of the final release for DHC output.

How does this compare to the predictability of initial data on other official series? Chart 3 plots the standard deviation of revisions and compares them to those of manufacturing output. Remember that the sectors are comparable in size in terms of GDP (both about 15% in 2004). The standard deviations are calculated over revisions of different horizons, and a similar real-time data base of manufacturing output was used. Naturally the longer the horizon the more volatile are revisions from the initial release, although that standard deviation should eventually converge. The chart shows that DHC data revisions are more volatile than manufacturing, whose revisions die down more quickly.

Chart 3: Standard deviations of revisions to official data



Source: ONS and own calculations.
Notes: Each standard deviation was calculated on a sample of no less than 80 observations.

This might be better understood when we consider how the official data on the sector is actually put together. In the ONS data the DHC sector corresponds to sections G and H in the UK Standard Industrial Classification (SIC). It combines the motor trades (including retail petrol sales), wholesale and commission trades, the retail goods trade and hotels and catering. The total sector covered around 15% per cent of GVA in 2004, with retailing being about 5.5%, wholesale about 4.5%; the motor trade about 2% and hotels and catering about 3%. As with other parts of the service sector, DHC has been increasing as a share of the UK economy, especially in terms of employment.

The official series for DHC output is compiled by the ONS as follows. The retail sector is measured by taking the retail sales turnover from the retail sales inquiry, and then deflating by Retail Price indices. Since 2002, wholesale and the motor vehicles sector outputs are estimated by a monthly survey, the Monthly Inquiry into the Distribution and Service Sector or MIDSS, which could then be combined with its pre-existing quarterly equivalent. In the case of motor vehicle, this improved on the previous method of combining data on new vehicle registrations, estimates of used cars turnover, and fuel deliveries to retail fuel centres from refineries. For example the survey should better capture other revenue activities such as servicing parts and repairs, it brought in direct information on used car sales, and better accounted for quality changes. These volumes are deflated by the RPI and the Corporate Service Price Index. Similarly for the wholesale sector, the new survey should improve on the previous method of extrapolating from the retails sales index and index of production proxies. For example, direct surveying should take better account of the wholesaling of imported goods. Where wholesale turnover is measured in value, it is deflated by elements of the Producer Price Index. The hotels and catering sector is mostly now captured by the MIDSS, with Corporate Service Price and Retail Price indices used to deflate turnover.

Clearly the adoption of this direct survey should have improved the measurement of the sector. But some difficulties will no doubt remain. These may be to do with the adjustment for input prices. The point here is that the output measure is meant to capture value-added, not gross output. The value of intermediate inputs into each industry have to be subtracted from the value of turnover and changes in stocks and own account capital formation added. But input prices, for example the price of imported clothing or of ICT related products can change rapidly, sometimes falling much faster than output prices. Ideally, the real value-added measure should adjust for this, but that requires more frequent measurement of input prices than is typically possible. Much of the DHC output is estimated by assuming that changes in net output are approximated by changes in gross output; that is, by assuming that margins are constant. There are good reasons to argue why that might not be too distortive — see Office for National Statistics (2006b) — but there might still be some source of error, especially in initial releases.

In summary, later releases of the official data on our variable of interest will contain more information than early releases. It takes time for official data to incorporate and process information. But we can expect that the measurement of this sector would improve with the recent adoption of a direct monthly survey. However even if measurement has improved, innovation in these sectors, associated with the use of information technology for example in the adoption of electronic point of sale systems may have thrown up new measurement challenges. See Moir and Dawson (1992) and Dawson (2004).

2.2 Other official data and non-official indicators

Under these circumstances, one might turn to data from other official and non-official sources. Considering first the other official data, there are potentially many series which one could use; remember that we are looking to extract only a marginal contribution from each. For example, data on the consumption of goods, household income, employment and investment of the DHC sector might all bring in relevant information here.

However there is one good reason to limit the number of other official series. Other official series are themselves subject to revision, but unlike for DHC output, a real-time data base is typically not available. I would be forced to use the latest estimate of the official series and would thus be mixing in information from different vintages. This would mean that the observations in our equation suffer from a heteroscedastic measurement error: the further back in the sample, the more precise would the series be. The econometric estimates might be affected — they may be erroneously giving the same weight to more imprecise recent early releases of official data as they would estimate for past mature data: introducing biases to the forecast.

So I have restricted myself to only two other official series. The first chosen candidate was real household disposable income, shown on Chart 4. It is hoped that household income data would bring in some information about household spending decisions. Of course short-term movements in income need not be spent: it can be saved, especially following the rapid financial innovation in the UK since the 1980s. All the same, to the extent that consumers are liquidity constrained, or that they perceive movements in disposable income as permanent, there might be some relationship to exploit.

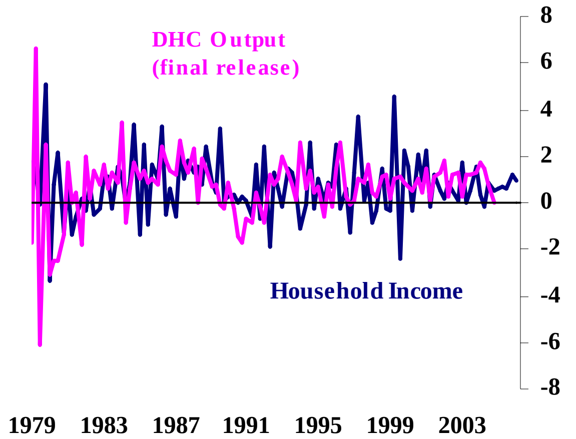
The second official series is the volume retail sales survey. See Chart 5. This survey is conducted by the ONS and contains very similar data to the early releases of DHC output. One difference is that it captures sales volumes, and not value-added and so weights the same raw information differently. Its advantage is that it is more timely than the initial DHC output data. This survey gets revised, but there is a real time data base available and I was able to use the series of initial estimates of retail sales.

I have also included one non-official survey indicator in the data set. The CBI distributive trades survey, shown in Chart 6, is a balance of respondents' from the total distribution sectors' answers to a qualitative question about the current volume of sales. Although it is timely, we would not necessarily expect this to be a very good measure of DHC output. For one thing, the CBI survey aims to capture the volume of sales, and like retail sales, it is not a value-added measure. And then, the question asks respondents to compare current business conditions with those a year ago. This might lead to a dynamic pattern in the residuals when I compare it against quarterly growth. And its sample of firms is smaller than that of the official data; it does not cover hotels and catering; and being more timely it may not be adjusted to the same quality as the official data[3].

Still, looking at the chart, the non-official series does show some common relationship with final release on DHC output. That is reassuring, although it is the marginal information we might want to extract.

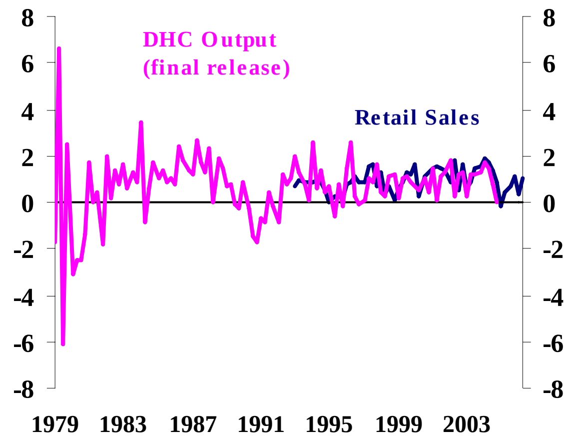
[3] I also experimented with including the CIPS Purchasing Managers Index survey for the service sector. The CIPS actually covers all other private service sector except for distribution, hotels and catering. It might still have mattered though, if there were shocks common to all service sector industries. However, after testing it proved to be insignificant.

Chart 4: DHC output and household income



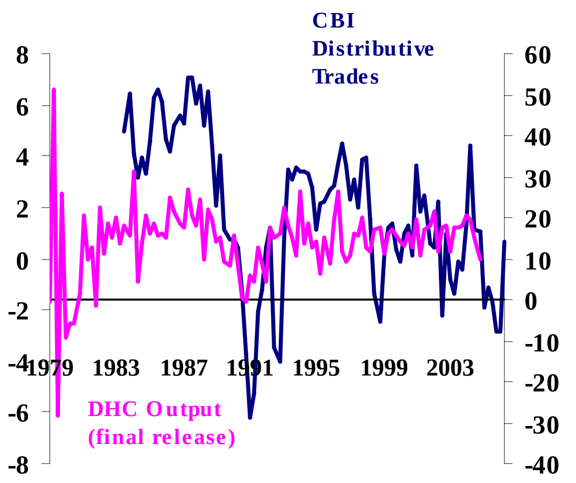
Source: ONS

Chart 5: DHC output and retail sales



Source: ONS

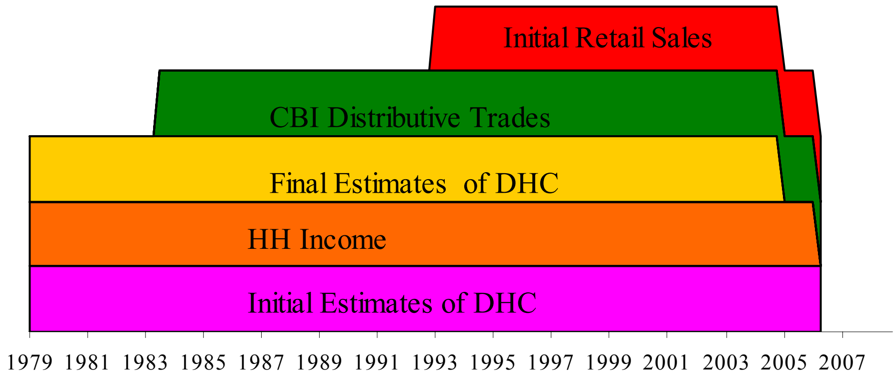
Chart 6: DHC output and CBI distributive trades



Source: ONS and CBI

The eclectic sources have left us with a very unbalanced data set. This is presented visually in Chart 7. Notice that the final release of the DHC output cuts off five quarters before the initial release (by definition). And the surveys although short in sample provide a potentially valuable extra quarter's information above the latest release of household income data. Initial releases for the current quarter are not available early on in that quarter, but it so happens that this paper was written towards the end of a quarter.

Chart 7: Sample lengths of the 2006 Q2 vintage data set



2.3 The structure for the truth

The one other source of information is the structure for true distribution sector output growth. In a model with no data uncertainty it would be sufficient to define an underlying economic process. But in a model of data measurement, we also need to define how that process links to the observed data. The structure is formed from these two parts.

True DHC output growth (g_t) is assumed to be first-order autoregressive,

$$g_t = \gamma_1 + \phi_{11} g_{t-1} + v_t + u_{1t} \tag{1}$$

with an economic shock which is itself autoregressive:

$$v_t = \phi_{22} v_{t-1} + u_{2t} \tag{2}$$

Both u_{1t} and u_{2t} are normally distributed mean-zero process with variances of σ_{u1}^2 and σ_{u2}^2 respectively. They are assumed be independent of each other, without much loss of generality. The equilibrium of the process for true DHC output implies a constant growth rate, $\gamma_1 (1-\phi_{11})^{-1}$, and that would be where DHC output growth would settle to if we were to forecast it far enough in the future for any current data information not to matter.

A second aspect of our definition is the relation between true DHC output and the observed data. Here I have taken the line that the truth we are interested in is exactly the final releases of official data, and that is measured by the data in the latest release that has been through at least five quarters of revisions. If f_t is that measure of final official data release for DHC output, identification is determined by

$$f_t = g_t \tag{3}$$

This choice is justified in the introduction. Returning to Chart 3, we an see that about two-thirds of the rise in the standard deviation of the revision has happened after five quarters. Equations 1, 2 and 3 define a simple structure. More complicated alternatives are certainly possible. For example one could allow for dynamics in equation 3. But there is a cost to introducing complications, in that the data may not be able to separately pick out the different parameters of more complicated mechanisms. What is important here is to recognise that equations 1, 2 and 3 identify the truth, and in this way help us to estimate it and also to explain it.

To make this model useful, the parameters γ_1, ϕ_{11}, ϕ_{22}, σ^2_{u1} and σ^2_{u2} all need to be estimated, and the initial conditions for the state variables (g_0 and v_0) also pinned down by a mixture of estimation and assumption. In principle this model might be estimated on the data for final release alone. But that would ignore the potentially useful role of initial data on DHC output and our other series.

3 The structure

Equations 1, 2 and 3 are set within an extended state-space model to bring in information from other data. The structure is as laid out in Harvey (1989) page 337, or in Cunningham, Jeffrey, Kapetanois and Labhard (2006). There are two unobserved state variables: true DHC output growth, g_t, and the autoregressive element of the economic shock to true growth, v_t. Following 1 and 3, the vector of states, $\alpha_t \equiv [g_t, v_t]'$, is determined by the process,

$$\alpha_t = \Phi \alpha_{t-1} + \Gamma + u_t \, , \tag{4}$$

with

$$\Phi \equiv \begin{bmatrix} \phi_{11} & 1 \\ 0 & \phi_{22} \end{bmatrix}; \tag{5}$$

$$\Gamma \equiv \begin{bmatrix} \gamma_1 \\ 0 \end{bmatrix}; \tag{6}$$

and the variance-covariance matrix of the state residuals assumed to be diagonal:

$$Q \equiv \begin{bmatrix} \sigma_{u1}^2 & 0 \\ 0 & \sigma_{u2}^2 \end{bmatrix}. \tag{7}$$

The vector of observed variables (y_t) brings together five data series — the final releases, the initial release (p_t) and the three other indicator variables:

$$y_t \equiv [f_t \quad p_t \quad y_{3t} \quad y_{4t} \quad y_{5t}]^T.$$

This is related to the unobserved state by

$$y_t = H\alpha_t + D_t + w_t \tag{8}$$

where the indicator measurement errors are

$$w_t \equiv [0 \quad w_{2t} \quad w_{3t} \quad w_{4t} \quad w_{5t}]^T \tag{9}$$

with a covariance matrix

$$R \equiv \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{w22}^2 & \sigma_{w23} & \sigma_{w24} & \sigma_{w25} \\ 0 & \sigma_{w23} & \sigma_{w33}^2 & \sigma_{w34} & \sigma_{w35} \\ 0 & \sigma_{w24} & \sigma_{w34} & \sigma_{w44}^2 & \sigma_{w33}^2 \\ 0 & \sigma_{w25} & \sigma_{w35} & \sigma_{w45} & \sigma_{w55}^2 \end{bmatrix}. \tag{10}$$

The other matrices are

$$H \equiv \begin{bmatrix} 1 & h_{12} & h_{13} & h_{14} & h_{15} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T, \tag{11}$$

and

$$D_t \equiv \begin{bmatrix} 0 & d_2 & d_3 & d_4 & d_5 \\ 0 & 0 & d_{8t} & 0 & 0 \end{bmatrix}^T. \tag{12}$$

The first line of the matrix equation 8 is just equation 3, which defines how the truth is related to the final release data. As there is no measurement error in this identity, the first row and column of R are empty. Consider now the four other measurement errors (w_{2t} , w_{3t} , w_{4t} , w_{5t}) in the context of equation 3. Each of these are the only measurement error in each indicator observation equation. But they can be correlated with each other, and the relationship between the true growth and the indicator need not be one-for-one. They are not correlated with the economic shock. For example our model of the initial release is given by the second row of equation 8,

$$p_t = h_{12} g_t + d_2 + w_{2t} , \tag{13}$$

where w_{2t} can be correlated with the other measurement errors only and h_{12} need not be one.

Note also that the matrices Φ, Γ, Q, H, and R are assumed to not vary over time: the economic structure is assumed to be time invariant. D_t does vary over time, but only insofar as d_{8t} is a shift dummy, shifting from zero to one at and beyond 1989Q1. This is to allow for a shift in the measurement process for the CBI indicator. I justify this later on.

But the number of different series in the data set does vary over time. How is the estimation method adapted to cope with this? As an illustration, consider the simple case that only one observation — the value of indicator 3 for a period $t = m1$ — is missing. I define a set of time-varying matrices, U_t of size L_t by 5 ($L_t = 1, \ldots, 5$ for $t = 1, \ldots, N$) such that:

$$U_{m1} = \begin{bmatrix} I_2 & 0 & 0 \\ 0 & 0 & I_3 \end{bmatrix}, U_{t \neq m1} = I_6 .$$

Then our measurement equation becomes

$$U_t y_t = U_t H \alpha_t + U_t D_t + U_t w_t . \tag{14}$$

Replacing 8 with 14 means that the Kalman filter narrows and broadens depending on whether fewer or more data series available. Our estimated model now has time-variation in the parameters of the measurement equation. That said, the time-variation that this creates is imposed and need not be estimated.

I have imposed strong assumptions. Some could be relaxed. For example it might be worth allowing for time-variation in the variance-covariance matrices Q and R to capture some heteroscedasticity. But having said that, the model does still allow for many interesting possibilities. For example, what would we expect to see if a particular indicator, say the initial release, were a very good predictor of the truth? In that case the econometric estimates should have a value of h_{12} close to one, d_2 close to zero — indicating no bias; and σ^2_{w22} close to zero — a small variance of indicator measurement error.

Later on I show how the system comprising 4 and 14 can be estimated using a maximum log-likelihood procedure. Given those estimates, the outputs we would also want are as follows:

- First the estimated values of the parameters, and their standard errors.
- Second, what would be the monetary policymakers' estimate of true DHC output growth, formed on the full set of available data. This corresponds to the estimated value of the state conditional on the whole data set and are called the fixed-interval smoothed values of the state. They are denoted by $\hat{\alpha}_{t|N}$, and come with a covariance matrix $\hat{P}_{t|N}$. N is the size of the full available data sample. With the smoothed values of the state, one can also calculate a set of smoothed fitted values for the indicators, $\hat{y}_{t|N}$. The smoothed residuals are then the difference between the observed indicator values, and these fitted values and are denoted by $e_{smooth,t} = y_t - \hat{y}_{t|N}$.
- Third, we would want also want a forecast of true distribution sector, at least for the next two quarters. That comes from a prediction of the state, denoted by $\hat{\alpha}_{N+k|N}$ where $k = 1, 2, ..., K$, and $N + K$ being the end of a forecast horizon. We can estimate a covariance associated with that prediction, $\hat{P}_{N+k|N}$.
- Fourth, we could look at the contribution of each data observation to the estimate of true DHC growth. Formally I derive the contribution of observation on indicator i at time t to the estimate of the true DHC growth at time s, denoted by $c_{i,t,1,s}$ such that

  $$\sum_{i=1}^{5}\sum_{t=0}^{N} c_{i,t,1,s} = \hat{g}_{s|N} .$$

  This is done by running the recursions above over a data set that includes only the value of that indicator at time t and with all other values are set at zero[4]. As the filter is linear the sum of these contributions should add up to the true DHC growth estimate, as described above, when we also take account of the constant. How do we treat the missing values in calculating these contributions? The natural solution is to impose the set of matrices U_t when we calculate the filtered, smoothed and predicted values of the states. This means I am interpolating any missing values in deriving our outputs, and there is therefore no contribution of a missing observation.

[4] A more computationally efficient alternative is to run the recursions described in Koopman and Harvey (2003).

- Fifth, there are the weights. The contribution of each data series at time t to the estimate of the state at time s is given as the value of that series at time t multiplied by its weight, $c_{i,t,1,s} = y_{it} w_{i,t,1,s}$ such that

$$\sum_{i=1}^{5} \sum_{t=0}^{N} y_{it} w_{i,t,1,s} = \hat{g}_{s|N} \ ,$$

thus breaking the contribution further down into the weight and the data outturn. In a Kalman filter with variable matrices, the weights vary over both the timing of the data, t, and the timing of the predicted DHC output, s. The weights for calculating the estimate of true DHC growth in the most recent quarters (s near to N) are likely to be of most interest.

- Finally, we can also calculate the information content of each indicator, which is related to its contribution in reducing the variance of the estimation error to the true growth measure.

4 How the model is estimated

A first step is to prepare the data. All observed series are demeaned and scaled to the final release (multiplied by the ratio of the standard deviation of the final release to the standard deviation of that data series) on a common sample. This restricts the range of parameters, and so improves the speed and accuracy of the optimisation. For example, de-meaning should push the values of the estimated constants in observations closer to zero. This is especially important because not all of our indicators are growth rates. The survey data are balance measures which report the percentage of respondents and have a very different scale.

Demeaning and scaling are not exact procedures though. When there is substantial uncertainty in estimating the sample mean and standard deviation of a series, these procedures will naturally be subject to uncertainty. So I have allowed for constants in the relationships which are not identities, even after demeaning, and then also tested for shift structural breaks.

The estimation method is (non-Bayesian) maximum likelihood[5]. The log-likelihood function is minimised over the choice of the parameter values by a numerical algorithm given some starting values for the parameters and the states. That process and the assumptions used are described in Appendix 1.

[5] Under Bayesian maximum likelihood, the distribution of the observed data would also conditional on priors about distributions of the initial state and about the distributions of the parameters.

I only note here that the starting values for the estimated parameters were taken to be fixed and not stochastic. The numbers chosen are presented along with the estimation results later. The starting values of the filtered states $\hat{\alpha}_{1|0}$ and its covariance $P_{1|0}$ were set at their the unconditional mean and covariance, $\left(I-\hat{\Phi}_0\right)^{-1}\hat{\Gamma}_0$ and $\hat{\Phi}_0\left(\overline{P}_0\right)\hat{\Phi}_0^{-1}+\hat{Q}_0$, respectively. This is valid if the state variables are stationary.

5 The parameter estimates

The estimates of the unrestricted parameters are given in Tables 2, 3 and 4.

Table 2: Estimates of the state equation	
$\hat{\Phi}$	
$\hat{\Phi}_{11}$	0.3698 (0.1026)** (0.5)
$\hat{\Phi}_{22}$	-0.6880 (0.0453)** (0)
$\hat{\Gamma}$	
$\hat{\gamma}_1$	0.3875 (0.1105)** (0.675 $*$ 0.5)
$\hat{Q}$	
$\hat{Q}_{11}$	0.8583 (0.1880)** (0.01)
$\hat{Q}_{12}$	
$\hat{Q}_{22}$	1.601 (0.1082)** (0.01)
Notes	The first figure in brackets is the estimated standard error.
	The second bracketed figure is the starting value used of that parameter.
	One star indicates 90% significance, two stars indicate 95% significance.

Table 3: The estimated covariance of the measurement equation

$$\hat{R}=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1.0678\,(0.1394)^{**}\,(0.01) & & & \\ 0 & 0.4356\,(0.1838)^{**}\,(0) & 2.1580\,(0.3229)^{**}\,(0.01) & & \\ 0 & 1.5465\,(0.4563)^{**}\,(0) & 1.5053\,(0.5663)\,(0) & 3.6686\,(1.2035)^{*}\,(0.01) & \\ 0 & 0.0004\,(0.1376)\,(0) & 0.0374\,(0.2980)\,(0) & 0.2664\,(0.3372)\,(0) & 2.3955\,(0.3210)^{**}\,(0.01) \end{bmatrix}$$

Table 4: Estimates of the measurement equation

$$\hat{H} = \begin{bmatrix} 1 & 0.6679\,(0.0687)^{**}\,(1) & 0.7707\,(0.1774)^{**}\,(1) & 1.4930\,(0.3291)^{**}\,(1) & 0.2543\,(0.0960)^{*}\,(1) \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T;$$

$$\hat{D} = \begin{bmatrix} 0 & -0.3466\,(0.1113)^{**}\,(0) & 0.1063\,(0.2392)\,(0) & -0.5258\,(0.3140)\,(0) & 0.5103\,(0.1510)^{**}\,(0) \\ 0 & 0 & -2.4921\,(0.3678)^{**}\,(0) & 0 & 0 \end{bmatrix}^T.$$

Taken as a whole, this comprises a reasonable set of estimates: most of the parameter estimates are significant; all fall within the range of what is plausible; and all have been updated from their initial values. Looking at them more closely, we can see that there are some interesting results.

- True DHC output growth is not very autoregressive. The coefficient on the lagged value is about 0.4.
- True DHC output growth is estimated to be quite volatile. The conditional standard error of economic shocks ($\left(\hat{Q}_{11}\right)^{0.5}$) is about 0.93 = $(0.86)^{0.5}$. That is larger than the mean of the final release of official data for that sector (0.68 %). The unconditional standard error is about $0.99\% = \left(\left(1+\left(\hat{\Phi}_{11}\right)^2\right) * \hat{Q}_{11}\right)^{0.5}$.
- The long-run growth rate of true DHC output is estimated to be $0.61 = \left(1-\hat{\Phi}_{11}\right)^{-1}\hat{\gamma}_1$. Our predictions of true DHC output growth into the quarters beyond which there is no data become quickly dominated by this estimate. But it is not very precise. The estimate might be sensitive to very early quarters in our data set, where there were some large falls in the sectors' output. See Chart 1. As a rough indication of the imprecision, the standard error on $\hat{\gamma}_1$ is 0.11, nearly half its estimated mean. I shall return to this later on.
- Initial releases are imperfectly related to true DHC growth. The estimate of the long-run coefficient linking the two, $\hat{h}_{21}$, at 0.67 is significantly below one. There is also a significant constant under-prediction, of about 0.35pp, as estimated in $\hat{d}_2$. And the unconditional standard error of its measurement error is estimated to be large, at 1.05%[6].
- The CBI distributive trades indicator has an upward bias which is not significant. It's long-run relation with the truth ($\hat{h}_{31}$) is less than one, but significant. And the unconditional

---

[6] The sample estimate of the standard deviation of the measurement error in the initial release $E[p_t-\alpha_t-E[p_t-\alpha_t]]^2$ is calculated as $\left(\left(1-\hat{h}_{21}\right)^2 * \hat{Q}_{11} * \left(1-\left(\hat{\Phi}_{11}\right)^2\right) * \left(1-\left(\hat{\Phi}_{22}\right)^2\right) + \hat{R}_{22}\right)^{0.5}$, and similarly for the other series.

standard deviation of its measurement error is larger than even that of the preliminary release at 1.47%. The shift dummy $\hat{d}_{8t}$ was found to be significant.

- The initial retail sales data is also significantly related with true DHC growth: see $\hat{h}_{41}$. The unconditional standard deviation of its measurement error is 1.94%, larger still. It features a downward bias ($\hat{d}_4 < 0$) but one that is not significant. Note that, just as we would expect, there is significant covariance between its idiosyncratic measurement errors and that of the initial data release: $\hat{\sigma}_{24} > 0$; the correlation coefficient is about 0.78. This suggests that some small part of errors in the initial ONS series reflect retail sales information.
- Households' income is also found to be weakly related to true DHC output, at least by these measures.

But a ranking of indicators in terms of the standard deviation of measurement error does not necessarily translate into a ranking of their contribution in estimating the truth. The marginal information in a particular data outturn brings in to predicting DHC output is a system property, and can differ even in sign from the estimated coefficients in the observation matrix, H . In particular it can be affected by error covariance. I shall derive full system measures of how much each indicator matters to complement these partial measures in sections 7 and 8.

6 Diagnostics

Charts 8 and 9 plot the one-step ahead residuals for the final and initial release. These residuals are all standardised.

Chart 8: The one-step ahead residuals for the final release

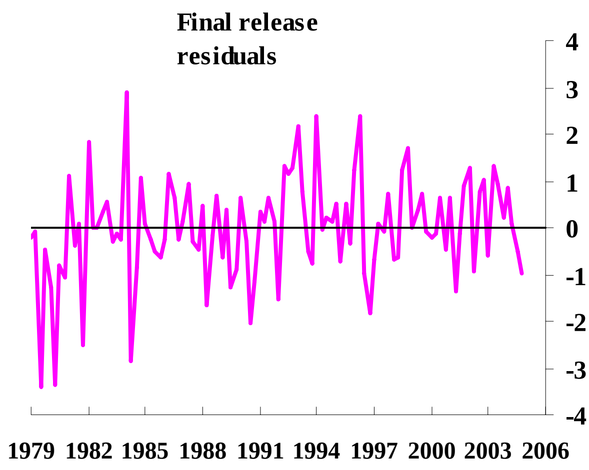
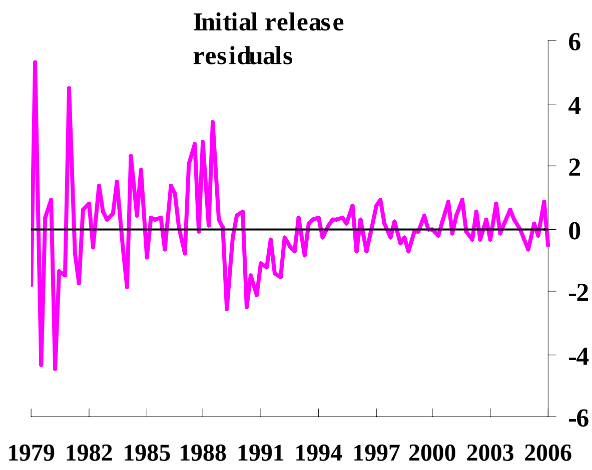


Chart 9: The one-step ahead residuals for the initial release



Problems with the model can be found by testing on these residuals and also those for the indicators. We can certainly see some heteroscedasticity from those charts. And, though the residuals are quite noisy, that might disguise a break of some sort. Looking at the CBI series in Chart 6, there seems like there was a structural break in the late 1980s. Indeed a shift dummy was introduced to deal with this after testing (see below).

The tests are for serial correlation, heteroscedasticity; non-normality and finally for structural breaks with an unknown timing for the one-step-ahead residuals. Table 5 presents the statistics with the probability value in brackets.

Table 5: Diagnostics for one-step ahead residuals on observation equations					
	Final	Initial	CBI DT	RS	HI
No of obs	104	110	92	54	109
Qstat ACF 1	5.563 (0.018)	2.050 (0.152)	39.479 (0.000)	5.125 (0.023)	16.652 (0.000)
Qstat ACF 4	8.511 (0.075)	7.557 (0.254)	62.791 (0.000)	13.901 (0.008)	19.663 (0.000)
Serial LM test 1	6.267 (0.012)	1.822 (0.177)	38.669 (0.000)	5.107 (0.024)	16.600 (0.000)
Serial LM test 4	4.700 (0.320)	4.393 (0.355)	38.435 (0.000)	12.004 (0.017)	19.188 (0.000)
Engle LM test 1	14.423 (0.000)	6.460 (0.011)	19.753 (0.000)	0.299 (0.585)	8.684 (0.003)
Engle LM test 4	47.9187 (0.000)	36.929 (0.000)	23.032 (0.000)	0.494 (0.974)	16.219 (0.003)
Jacque Bera	189.158 (0.000)	76.426 (0.000)	10.930 (0.000)	1.410 (0.494)	6.743 (0.034)
Kolmogorov	0.080 (0.504)	0.079 (0.392)	0.172 (0.07)	0.175 (0.065)	0.130 (0.050)
Breakpoint test	3.699 (0.425)	6.289 (0.125)	2.21 (0.659)	5.52 (0.159)	2.940 (0.523)
Bootstrap 95% crit value	9.793	8.186	9.090	8.474	8.660
Dates of three largest breaks	1982Q4, 1983Q1,1983Q2	1988Q4, 1989Q1, 1988Q3	1984Q1, 1984Q2, 1984Q4	1982Q4, 1985Q1, 1985Q2	1982Q4, 1983Q1, 1983Q3
Statistic with p-value in brackets					

Serial correlation is tested with a Ljung-Box Q-statistic for one and up to four lags, and also with Breusch-Godfrey LM tests for one and four lags (both Chi-squared with one degree of freedom). The results indicate that the residuals from the CBI, retail sales, household income all suffer from some first-order serial correlation, or some mis-specified dynamics.

There is some first-order correlation in the final release. But that is not at all worrying. The final release residuals should capture the economic shocks to the model. So in some sense we should expect serial correlation here; the model takes account of that. And reassuringly, the initial release residuals are found to be not serially correlated.

The Engle LM test indicates heteroscedasticity in all residuals of order one, and up to four in all of the residuals except retail sales. This may be expected of most UK time series over this sample; the 1980s were simply more volatile than later decades. Another potential cause is that there may be large shifts in the variance of measurement error. Garratt, Koop and Vahey (2006) suggest that a failure to account for these large shifts with regime switching might mislead model predictions. But they test a regime switching model against a linear, fixed coefficient OLS version; a state-space model even with constant variances such as this may imply less predictive failure.

The Jarque-Bera test (Chi-squared with two degrees of freedom) indicates a rejection of the null of normality, in the sense of kurtosis and skewness, in all but the Retail Sales residual series at a 5% level. The other two other series also come close to rejecting that null. The Kolmogorov-Smirnoff statistics reported in the next column tests if the residuals come from a normal distribution, and all series pass at a 5% level but only just. Yet a failure of normality and heteroscedasticity is not disastrous. The Kalman filter is the unbiased estimator with minimum variance if the residuals are gaussian. But if they are not, the Kalman filter is still the linear minimum variance estimator (though it might be bettered by a non-linear estimator).

The structural break test allows for breaks of an unknown timing. I used the Supremum test of Quandt (1960) and Andrews (1993). A series of Wald tests for permanent shift structural breaks were carried out, within a trimmed sample[7]. The test statistic is the maximum of this series. The critical values are derived by a wild bootstrap method of O'Reilly and Whelan (2005), as that should improve on the asymptotic values of Andrews (1993). The Wald statistic is reported in the column, with the date of the three largest breaks below that, and then the bootstrap 5% critical value itself.

The statistics indicate that neither the final release residuals nor the initial release residuals suffer from significant structural breaks, although there is more likelihood of a break in the initial

[7] A trimming parameter of 0.15 is applied. See O'Reilly and Whelan (2005) for an explanation.

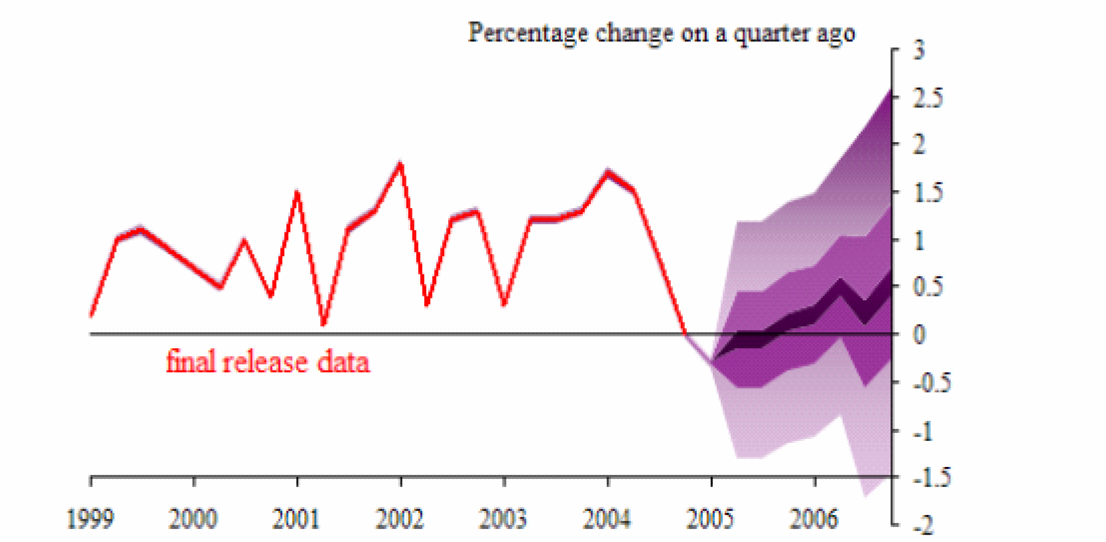
release residual in 1989. That is reassuring, as the final release data is important in identification. Note the important 2002 changes in methodology in 2002 (see Section 2.1) are not indicated as a structural break in the official data residuals. Perhaps more data would be needed for these changes to become visible. It could also be that the effect of these changes is more of a reduction in variance rather than a step shift. None of the other indicator measurements are found to suffer from breaks. In the case of the CBI series, this is because we have already adjusted for a break with a shift dummy; without that dummy the residuals fail the break test unequivocally.

Of course these tests might be improved upon. For example it is difficult to identify structural breaks of unknown timing independently of estimating the model's parameters. As the one is conditional on the other, it may be important to do the two together. This matters especially in a Kalman filter framework. This has been tackled in a Bayesian framework, by Giordani, Kohn, and Van Dijk (2005). Still, these diagnostics do reveal some problems. Heteroscedasticity is ubiquitous and serial correlation and non-normality is present in many of the indicator residuals. But these concerns do not seem grave enough to undermine the main findings.

7 The forecast, the smoothed estimates, and the contributions

Chart 10 plots the estimates of the distribution of true DHC growth as a fanchart. The central band of 10% width tracks the final release data quite closely where that data is available. In part that is by construction, as true distribution sector growth is defined as the long run of the final release series. But it also matters that a characteristic of the final release series is that it has short-lived dynamics and a small measurement error and so keeps very close to this long run.

Chart 10: Estimated distribution of true DHC growth

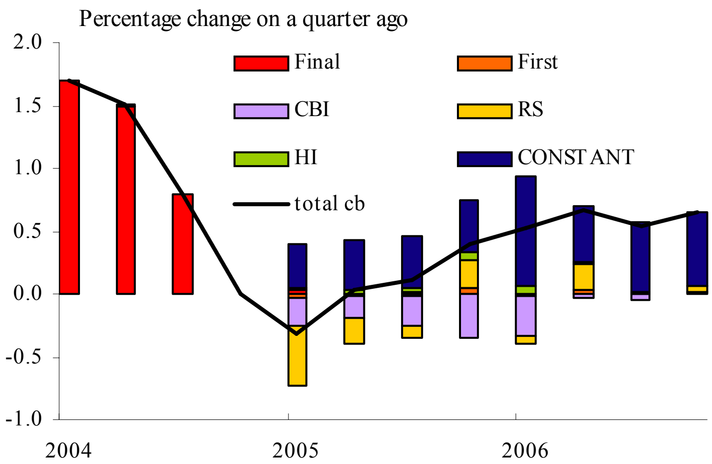


Note: Bands are of 10% width, 50% width and 90% widths about the mean.

Moving beyond the reach of the final release data, the chart confirms the pattern of a slowdown in early 2005 and a short interrupted recovery towards end of 2005, and another recovery in

2006Q2. Conditional on the available data set, the model suggests that the recovery will continue. But the standard error bands around the Q2 values are large: true DHC growth is estimated to lie within the range of -0.8% to 1.8% with a 90% probability.
Chart 11 breaks down the central estimate of true DHC growth into the contributions of the different data series. It includes what I would consider to be the time span of interest: for a data set put together in 2006 Q2, that is 2004Q4 to 2006Q4.

Chart 11: Estimated contributions to the estimate of mean true DHC growth

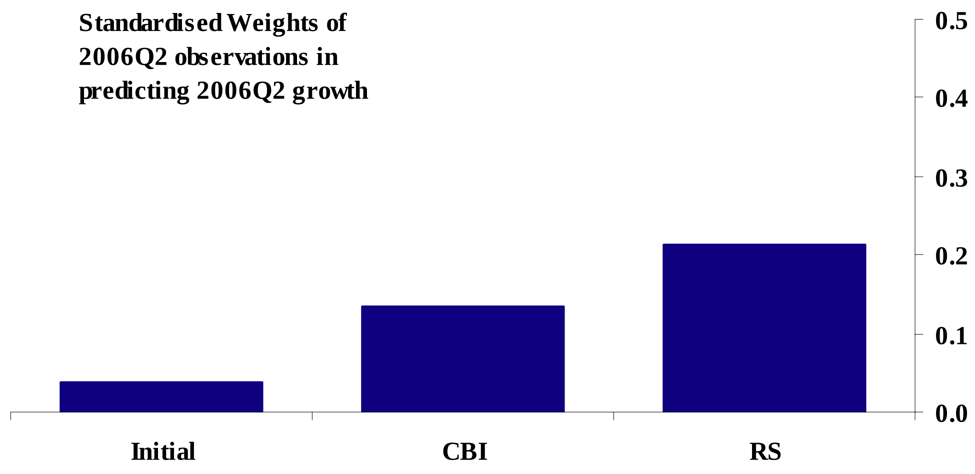


Where there is final release data (up to 2004Q4), that dominates in determining the estimating the mean of DHC output growth. Where there is no data, the estimated constant growth rate matters most. The dynamics of the model are short-lived. Hence the forecast of DHC growth for the last two quarters, 2006Q3 and 2006Q4, reflects little of the observed data and is dominated by the estimate of the constant γ_1.

In between these two time frames, the other indicators do play an important role. The contributions of the CBI Distributive Trades data and the initial retail sales series have acted to pull down on the mean estimates of DHC growth over the first three quarters of 2005. The retail sales data contribution turns positive in the end of 2005. It is striking that the initial data release does not matter much, and has an erratic contribution. I return to that below.

To understand those contributions better, Chart 12 plots those weights for the mean estimate of true DHC growth in the last period for which data is available; in our case that was 2006Q2. An initial release of official data for DHC growth for that period was available, although this would not be true early on within the quarter.

Chart 12: Standardised weights to mean forecast of true DHC growth



There is no final release data available for 2006Q2, nor household income data. In their absence the CBI data and the retail sales series matter the most. We see that the initial data, that is in this case available up to 2006Q2, earns only a small weight in predicting distribution sector output growth. Its weak contribution in the previous chart was not because the outturns came out close to its mean.

8 Information gain of the indicators

In the last section, I measured how important (or not important) our different data sources were in terms of contributions or weights to a central estimate. But we might be interested in how much each indicator contributes to improving the accuracy of the Kalman filter, hence as a variance decomposition.

This can be approached from the point of view of information theory. The idea is to estimate how much expected uncertainty is lowered by the information provided by each indicator. Tinsley, Spindt and Friar (1980) and Coenen, Levin and Wieland (2005) show how assuming that errors are normally distributed, that measure of the information gain of indicator i is related to the root mean squared error of the estimate of the true state when all indicators are used as a percentage of when all but the indicator i is used. This is also explained in Appendix 2.

Here I decompose the variance at two periods of time, at 2005Q1, two quarters beyond where the last final estimate was available and 2006Q2, the end of all data in the sample. Table 6 reports the results.

The calculations in Table 6 confirm that the initial release data provides less information to the estimation of true DHC output growth than do the other indicators. The most important uncertainty reduction is given by the CBI and the retail sales series. That said, the information gains from those series are arguably not large either, relative to the estimated constant in the estimates of the state variable.

In summary, I have three main findings. First the model indicates that there was a slowdown in 2005, but is less sure about the subsequent recovery (at least from the data set I had available at the time). Second the initial release receives a low weight. And third, the estimated constant plays an important role in the forecasts of the model, especially beyond two quarters ahead. I now turn to discuss these two last findings in more depth.

Table 6: Contribution of indicators to information content

with…	RMSE of estimate for 2005Q1	% RMSE reduction	RMSE of forecast for 2006Q2	% RMSE reduction
...all indicators	0.7640		0.7969	
...all but final	0.7828	100*(0.7828/07640-1)=2.43	0.7970	0.01
...all but initial	0.7646	0.08	0.7976	0.08
...all but CBI	0.7857	2.80	0.8212	3.00
...all but RS	0.7981	4.37	0.8351	4.67
...all but HI	0.7673	0.42	0.8005	0.45

9 Identification, multicollinearity and reliability

Why is the contribution of the initial release smaller than that of the surveys? One reason could be that the initial release is an estimate which processes only the information that the ONS have at that stage on the final release series. It is not a forecast of the final release, which would combine that information with a model-based estimate of the revision process. The surveys instead may aim more directly at the final release concept.

The retail sales data seems to be strongly related to the final release series. If this coefficient were poorly estimated, it would be overstating the importance of the retail sales series, and consequently undermining the initial release series. The data series all aim at a very similar concept. The model may not be set up well enough to separately identify different parameters, especially the error variances. For a combination of these reasons, the parameter estimates could be multicollinear. It is commonly argued that multicollinearity is not a serious problem. But we know that even if OLS estimates can still be unbiased and of minimum variance under multicollinearity, they can still be unreliable, in the sense that the estimated coefficients are highly sensitive to new observations. The symptoms are instability in estimates as new observations are added; and large but offsetting parameters between two collinear indicators in the same equation (Watson, 1983).

We can test for this within the state space representation. Following Burmeister and Wall (1982), we can look to see if the Hessian matrix of parameter estimates is nearly singular, indicating that there are some linear combination of quite different parameter estimates which would as well be supported by the data, or relatedly if the correlation between different parameter estimates is close to unity in absolute value.

It turns out that the two parameter estimates that are most correlated are d_4 and h_{14}, the constant and the slope coefficient in the observation equation for retail sales data. So there may be some problems there, presumably reflecting that the estimated coefficient between retail sales and the final release data (a rather high 1.5) is sensitive to the estimate of the mean of the retail sales series. But the estimated correlation coefficient between the two parameter estimates is, at 0.73, quite far from unity (Burmeister and Wall worried about values great than 0.996 in absolute value). Nevertheless, given our earlier concerns about the estimated value of h_{14} — it essentially implies a low weight of the initial release data — this does at least indicate that there could be some problem.

To test for sensitivity I substituted the estimate of this coefficient ($\hat{h}_{14}$ in Table 4) with a value one standard deviation smaller, thus shifting it from 1.49 to 1.17. In Table 7, I compare the estimate of the Kalman gain weights (see appendix 1) under the two assumptions.

Table 7: Sensitivity of the Kalman Filter weights for 2006Q1

	with base-line coeff	with smaller coeff on retail sales
Initial Release	0.011	0.064
CBI DT	0.031	0.038
RS	0.045	0.012
HI	0.011	0.016

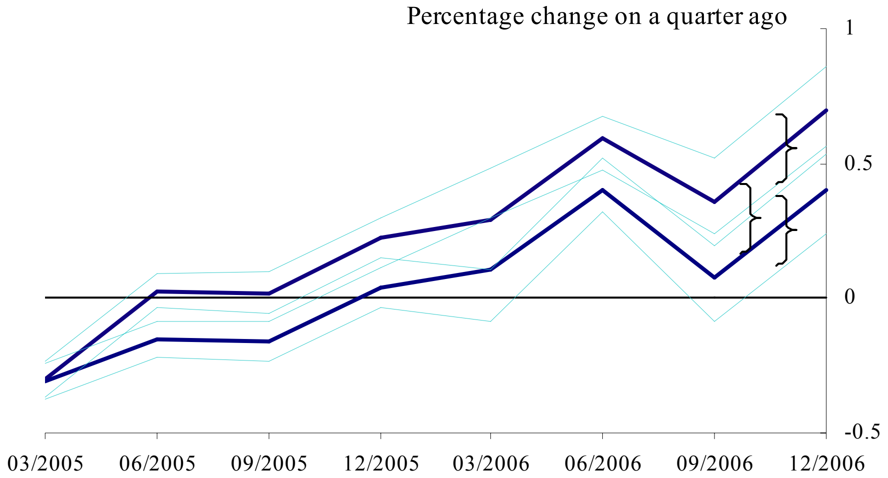
We can see that the weight on the initial release data rises nearly sixfold, from 0.011 to 0.064. So while one can accept the explanation that the initial release is not intended to capture the final estimate whereas the surveys are, one should also be wary that the low weight of the initial release may be due to an exaggerated role for the retail sales series. We can conclude that multicollinearity and weak identification pose no other problems here.

10 Robustness to uncertainty in the estimate of the long-run growth rate

The estimated constant plays an important role in our quantification. But that estimate may not be particularly well determined, especially within the ranges of accuracy that monetary policymakers need to work with. Yet the typical depictions of the uncertainty in this problem, such as the fanchart in Section 7, do not allow for uncertainty in the estimates of this or any other parameter. Another robustness check is needed.

Chart 13 plots three different 10% bands for true DHC growth. The central 10% band is for the model solved under our benchmark estimate of γ_1. The upper 10% band is for true growth solved with an estimate of γ_1 that is one standard deviation higher (about 0.1), and the lower 10% band is for an estimate of γ_1 one standard deviation lower.

Chart 13: DHC output: robustness



We can see that forecasts of true DHC growth are sensitive to uncertainty in this parameter. For example, as I shift the estimate of the parameter γ_1 by one standard deviation, the 10% bands for 2006Q2 are shifted up by about 0.17pp. That is because true distribution sector growth is estimated to revert to its mean quickly. There are two main implications. First this reminds us that this model is not that well suited for describing the next quarter beyond which there is no data (and certainly not beyond that) as essentially the forecast is the historical estimate of the average growth rate[8]. This is not a general result: forecasts from this type of model might be more robust for other series that are more backward-looking and revert more slowly to their historical mean. Second the estimates of the model are sensitive to shift structural breaks.

[8] This could be tested by comparing the root mean squared errors of out of sample forecast over different horizons.

11 Conclusion

In this paper, I have used a case study of distribution, hotels and catering output in order to explore the monetary policy data uncertainty problem. I defined a model of true output growth and estimated a distribution for the series using a mix of official and non-official data on the sector. The results confirm with some decent significance that the sector did undergo some slowdown in 2005 and that a recovery has started in early 2006. But they are less revealing about whether that recovery will last: the estimates for 2006 Q3 are uncertain, and quite dependent on an estimate for the average long-run growth rate, rather than on the new data releases.

That the estimated value of the state variable is dependent on the estimated growth rate reflects that DHC growth rate reverts to its historical mean fairly quickly. This has two important implications. First there is a role for off-model information in establishing the prior estimate of this growth rate. And second, I have shown that the model's predictions are quite sensitive to shift structural breaks. Testing for shift structural breaks in both the data measurement and in the economic process are therefore very important.

Another result is that the model places less weight on the initial release data than it does on the surveys. That could be because the initial release data is prepared by the ONS as a building block towards the final release rather than an estimate of it. And for this reason I showed that there is a danger that it could be due to a erroneous high weight on the retail sales data, which contains the same early available raw information.

Perhaps this model has been most successful in revealing how uncertainty is inherent in the monetary policy data problem. Even when data from different sources are used to inform, there is still substantial conditionality on particular assumptions. This would imply that there is a major role for discussing uncertainty when policy decisions are implemented and explained, but also testing for robustness. This is what I have tried to do.

Appendix 1: The Kalman Filter

Although the exposition of the Kalman filter is standard, it is worth repeating here so as to clarify notation. Given assumptions that the residuals are Gaussian, the Kalman filter produces the minimum mean squared linear estimator of the state vector α_{t+1} using the set of observations for time $t, Y_t = [y_1, ..., y_t]$. Let us call that estimate $\hat{\alpha}_{t+1|t}$ and its associated covariance $\hat{P}_{t+1|t}$. Estimates conditional on the whole sample of data are given recursively, taking t from $t = 1$ to N, as follows. I begin with a set of estimates for the parameter matrices $\hat{\Phi}, \hat{\Gamma}, \hat{H}, \hat{D}, \hat{Q},$ and $\hat{R}$, and a set of initial values for the state vector and its variance,

$\hat{\alpha}_{1|0}$ and $\hat{P}_{1|0}$.

Then the recursion runs from $t = 1$ to $t = N$ over

$$\hat{\alpha}_{t+1|t} = \left(\hat{\Phi} - K_t U_t \hat{H}\right)\hat{\alpha}_{t|t-1} + \hat{K}_t y_t + \left(\hat{\Gamma} - K_t \hat{D}_t\right) \tag{15}$$

where the gain matrix is given by,

$$\hat{K}_t = \hat{M}_t \hat{F}_t^{-1}, \tag{16}$$

and the covariance matrix of one-step estimation error is given by the solution to the Riccati equation,

$$\hat{P}_{t+1|t} = \hat{\Phi}\left(\hat{P}_{t|t-1}\right)\hat{\Phi}^{-1} + \hat{Q} - \hat{M}_t\left(\hat{M}_t\right)^T,$$

with

$$\hat{M}_t = \hat{\Phi}\hat{P}_{t|t-1}\left(U_t\hat{H}\right)^T,$$

and the covariance matrix of the one-step-ahead prediction errors in the observation data is estimated as

$$\hat{F}_t = \left(U_t\hat{H}\right)P_{t|t-1}\left(U_t\hat{H}\right)^T + U_t\hat{R}_t\left(U_t\right)^T.$$

The predicted estimate of the states, k periods ahead from the end of the data set $(k = 1, ..., K)$, is given by

$$\hat{\alpha}_{N+k|N} = \hat{\Phi}^k \hat{\alpha}_{N|N} + \sum_{j=1}^{k} \hat{\Phi}^j \Gamma$$

with a covariance matrix

$$\hat{P}_{N+k|N} = \hat{\Phi}^k \hat{P}_{N|N} \left(\hat{\Phi}^k\right)^T + \sum_{j=1}^{k} \hat{\Phi}^j Q \left(\hat{\Phi}^j\right)^T .$$

The fixed-interval smoothed estimates are instead given by working backwards:

$$\hat{\alpha}_{t|T} = \hat{\Phi}^{-1}\left(\hat{\alpha}_{t+1|t} - \hat{\Gamma}\right) + \hat{V}_t\left(\hat{\alpha}_{t+1|T} - \hat{\alpha}_{t+1|t}\right), \tag{17}$$

for $t = N-1, N-2, ..., 1$ with the associated covariance matrix as

$$\hat{P}_{t|N} = \hat{P}_{t|t} + \hat{V}_t\left(\hat{P}_{t+1|N} - \hat{P}_{t+1|t}\right)\left(\hat{V}_t\right)^T , \tag{18}$$

given the matrices,

$$\hat{V}_t = \hat{P}_{t|t} + \left(\hat{\Phi}\right)^T \left(\hat{P}_{t+1|t}\right)^{-1} .$$

the one-step ahead predictions for the observed data are given by

$$U_t \hat{y}_{t|t-1} = U_t \hat{H} \hat{\alpha}_{t|t-1} + U_t \hat{D};$$

The smoothed predictions follow from

$$U_t \hat{y}_{t|T} = U_t \hat{H} \hat{\alpha}_{t|T} + U_t \hat{D};$$

and the out-of-sample forecasts of the observed data are

$$\hat{y}_{t+k|N} = \hat{H} \hat{\alpha}_{t+k|N} + \hat{D}, \text{ for } k = 1, ..., K.$$

The covariance matrix of the smoothed prediction errors in the observation data is given by

$$\hat{G}_t = \left(U_t \hat{H}\right) P_{t|T} \left(U_t \hat{H}\right)^T + U_t \hat{R} (U_t)^T .$$

The log-likelihood function for the non-Bayesian problem is given as

$$\log L = -\frac{5N}{2} \log 2\pi - \frac{N}{2} \sum_{t=1}^{N} \log |F_t| - \frac{1}{2} \sum_{t=1}^{N} \left(U_t y_t - U_t \hat{y}_{t|t-1}\right)^T (F_t)^{-1} \left(U_t y_t - U_t \hat{y}_{t|t-1}\right).$$

The estimates of the parameters can be derived by numerically searching for values which maximise this log-likelihood. The numerical algorithm was chosen to be a Broyden-Fletcher-Goldfarb-Shanno algorithm. The principal idea of the method is to construct an approximate Hessian matrix of second derivatives of the function to be minimized, by analyzing successive gradient vectors, and this is much quicker than Newton methods. The initial conditions for the state are the starting values of the filtered states $\alpha_{1|0}$ and its covariance $P_{1|0}$ were taken to be the estimated unconditional mean and covariance,

$$\left(I - \hat{\Phi}_0\right)^{-1} \hat{\Gamma}_0$$

and

$$\hat{\Phi}_0 \left(\overline{P}_0\right) \hat{\Phi}_0^{-1} + \hat{Q}_0$$

respectively.

Appendix 2: Information theory

We can begin with Shannon's measure of uncertainty, or entropy, in a probability distribution. This is minus the expected value of the log of the probability distribution[9]:

$$J(x_{1t}) = E\left[-\ln\left(f(x_{1t})\right)\right];$$

$$= -\int f(x_{1t}) \ln f(x_{1t})\, dx_{1t}$$

9 Let us assume that one message is chosen from a finite set of equally likely (uniformly distributed) alternatives. Then the number of possible choices or any monotonic function of this number can be regarded as a measure of information. The logarithmic function is often judged to be the most natural function.

The joint entropy measures how much entropy is contained in a joint system of two random variables:

$$J(x_{1t}, x_{2t}) = E[-\ln(f(x_{1t}, x_{2t}))];$$

$$= -\int f(x_{1t})\, f(x_{2t}) \ln(f(x_{1t}, x_{2t}))\, dx_{1t} dx_{2t}.$$

The conditional entropy measures how much entropy a random variable x_{1t} has remaining if we have already learned completely the value of a second random variable x_{2t}:

$$J(x_{1t} \mid x_{2t}) = J(x_{1t}, x_{2t}) - J(x_{1t}).$$

Our measure of information content is then the reduction in uncertainty due to one indicator as the conditional entropy using the information set without that indicator minus the conditional entropy using the full information set:

$$\begin{aligned} I\left(\hat{\alpha}_{1t|N} \middle| Y_N^{-j}\right) &= J\left(\hat{\alpha}_{1t|N} \middle| Y_N^{-j}\right) - J\left(\hat{\alpha}_{1t|N} \middle| Y_N\right) \\ &= \left(J\left(\hat{\alpha}_{1t|N}\right) - J\left(\hat{\alpha}_{1t|N} \middle| Y_N\right)\right) - \left(J\left(\hat{\alpha}_{1t|N}\right) - J\left(\hat{\alpha}_{1t|N} \middle| Y_N^{-j}\right)\right). \end{aligned}$$

The joint distribution of the true state and the indicators, given by

$$v_t \sim N\begin{pmatrix} E\left[\hat{\alpha}_{1t|N}\right] & \\ & ,\Omega_t \\ HE\left[\hat{\alpha}_{1t|N}\right] + D & \end{pmatrix}$$

with

$$v_t = \left[\hat{\alpha}_{1t|N}, Y_N\right]^T$$

and

$$\Omega_t = \begin{bmatrix} E\left[\left(\hat{\alpha}_{1t|N}\right)^2\right] & U_t HE\left[\left(\hat{\alpha}_{1t|N}\right)^2\right] \\ U_t HE\left[\left(\hat{\alpha}_{1t|N}\right)^2\right] & U_t HE\left[\left(\hat{\alpha}_{1t|N}\right)^2\right](HU_t)^T + U_t R(U_t)^T \end{bmatrix}.$$

Tinsley, Spindt and Friar (1980) show that $I\left(\hat{\alpha}_{1t|N} \middle| Y_N\right)$ can be given by the log distance between the determinants of the marginal and conditional distributions of the estimates of the first state variable:

$$I\left(\hat{\alpha}_{1t|N} \middle| Y_N\right) = \frac{1}{2}\ln\left(\frac{\left|E\left[(\hat{\alpha}_{1t})^2\right]\right|}{\left|E\left[\left(\hat{\alpha}_{1t|N}\right)^2 \mid Y_N\right]\right|}\right);$$

$$= \frac{1}{2}\ln\left(\frac{\left|E\left[(\hat{\alpha}_{1t})^2\right]\right|}{\left|\hat{P}_{11,t|N}\right|}\right).$$

Similarly the contribution of each individual series can be given by the log distance between the determinants of the conditional distributions of the estimates of the state variables with and without that series:

$$I\left(\hat{\alpha}_{1t} \middle| Y_N^{-j}\right) = \ln\left(\frac{\left|\hat{P}_{11,t|-j,N}\right|^{0.5}}{\left|\hat{P}_{11,t|N}\right|^{0.5}}\right).$$

This is just the log difference between the RMSE of the estimate of the state without the indicator variable and with the full indicator set. Note that we have focussed on the effect of uncertainty only the first state variable, naturally. Less justifiably, we have ignored the role of estimation uncertainty in our measure.

References

Anderson, B D O and J B Moore (1994), `Optimal Filtering', Dover Publications, New York.

Burmeister, E and K D Wall (1982), `Kalman filtering estimation of unobserved rational expectations with an application to the German hyperinflation', Journal of Econometrics, 1982, vol. 20, issue 2, pages 255-284

Coenen, G, A Levin, and V Wieland (2005), `Data uncertainty and the role of money as an information variable for monetary policy', European Economic Review, Elsevier, vol. 49(4), pages 975-1006, May.

Clements, M P and D F Hendry (1999), `Forecasting Non-Stationary Economic Time Series', MIT Press, Cambridge, MA.

Cunningham A , C Jeffery, G Kapetanios and V Labhard, (2006), `A state space approach to the policymaker's data uncertainty problem'.Paper presented at the 38th annual Money Macro and Finance Conference 13 September 2006, available from http://ideas.repec.org/p/mmf/mmfc06/168.html'

Dawson, J (2004), `Retail change in Britain during 30 years: The Strategic Use of Economies of Scale and Scope', available at www.retaildawson.com/papers/Strategies.pdf

Garratt, T, Koop, G and S Vahey (2006), `Forecasting Substantial Data Revisions in the Presence of Model Uncertainty', University of Birkbeck Working Papers in Economics and Finance no 0617.

Giordani, P , R Kohn, and D J C Van Dijk (2005), `A Unified Approach to Nonlinearity, Structural Change, and Outliers', Econometric Institute Report EI 2005-09.

Harvey, A C (1989), `Forecasting, Structural Time Series Models and the Kalman Filter', Cambridge University Press, United Kingdom.

Koopman, S J and A C Harvey (2003), `Computing Observation Weights for Signal Extraction and Filtering', Journal of Economic Dynamics and Control, 2003, Volume 27, Issue 7, pages 1317-1333.

Kuttner, K N (1994), `Estimating Potential Output as a Latent Variable', Journal of Business and Economic Statistics, vol. 12, issue 3, pages 361-368.

Moir, C and J Dawson (1992), `Distribution: Reviews of the United Kingdom Statistical Sources', vol XXIX, Chapman and Hall.

O'Reilly, G and K Whelan (2005), `Testing parameter instability: a wild bootstrap approach', Central Bank of Ireland and Financial Services Authority, Research Technical Paper, No 8/RT/05.

Office for National Statistics (2006a), `Quarterly National Accounts Final Releases'.

Office for National Statistics (2006b), `Methodology of the Experimental Monthly Index of Services', http://www.statistics.gov.uk/iosmethodology/downloads/Whole_Report.pdf.

Oulton N and S Srinivasan (2004), `Productivity growth in UK industries, 1970-2000: structural change and the role of ICT', Bank of England working papers 259, Bank of England.

Quandt, R.(1960), Tests of the hypothesis that a linear regression system obeys two separate regimes, Journal of the American Statistical Association vol 55, pages 324-330.

Terceiro, J (1990), `Estimation of Dynamic Econometric Models with Errors in Variables', Berlin: Springer-Verlag.

Tinsley, P A, P A Spindt and M E Friar (1980), `Indicator and filter attributes of monetary aggregates: a nit-picking case for disaggregation', Journal of Econometrics, Elsevier, vol. 14(1), pages 61-91.

Watson, P K (1983), `Kalman Filtering as an Alternative to Ordinary Least Squares-Some Theoretical Considerations and Empirical Results', Empirical Economics, Springer, vol. 8(2), pages 71-85.