CARDIFF BUSINESS SCHOOL
WORKING PAPER SERIES

Cardiff Economics
Working Papers

Patrick Minford, Konstantinos Theodoridis
and David Meenagh

Testing a model of the UK by the method of indirect inference

E2007/2

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

ISSN 1749-6101
Updated April 2008

Testing a model of the UK by the method of indirect inference*

Patrick Minford[†‡§] Konstantinos Theodoridis[¶] David Meenagh[‖]

March 30, 2008

Abstract

We use the method of indirect inference to test a full open economy model of the UK that has been in forecasting use for three decades. The test establishes, using a Wald statistic, whether the parameters of a time-series representation estimated on the actual data lie within some confidence interval of the model-implied distribution. Various forms of time-series representations that could deal with the UK's various changes of monetary regime are tried; two are retained as adequate. The model is rejected under one but marginally accepted under the other, suggesting that with some modifications it could achieve general acceptability and that the testing method is worth investigating further.

JEL Classification: *C12, C32*

Keywords: *Bootstrap, Model Evaluation, Non-Linear Time Series Models, Indirect inference, open economy models, UK models*

*We are grateful for useful comments to James Davidson, Bernard Pearson and Simon Wren-Lewis; also to Mike Wickens for discussions over a long period. The paper was written before Konstantinos Theodoridis joined the Bank of England. The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England.

†Cardiff Business School, Cardiff University

‡CEPR

§Corresponding Author: Cardiff Business School, Colum Drive, Cardiff, CF10 3EU, Tel: +44 (0)29 2087 5728, Fax: +44 (0)29 2087 5728, E-mail: MinfordP@cardiff.ac.uk

¶Bank of England

‖Cardiff Business School, Cardiff University

The aim of this paper is to propose a new test of a model's dynamic properties and to put it to trial in an extended example on a macro model that has been in longstanding use, with the ultimate objective of evaluating this proposed procedure. Thus there are two main aspects to the paper. First, we set out this test of a model's dynamic properties, which uses the method of indirect inference, hitherto not used for this purpose but in our view well adapted for it. Second, we apply it to a model of the UK that has been in practical use for nearly three decades; because it can make claims both empirically to be close to the data and theoretically to approximate in key ways to recent DSGE models, it seems to be an interesting vehicle for giving the proposed test a trial run.

1 Related literature

Much recent work has been concerned with confronting structural macro models with the 'dynamic facts of the data'. These facts are collated by methods that are not related to the particular structural model in hand, though they should not of course be inconsistent with it; and so in principle they can be used to compare rival models with quite different theoretical structures. This sets the confrontation in these methods apart from the testing carried out in the estimation process. There a model is estimated and tested subject to the overall structure of the model, so that whereas features of the model may be tested and rejected, the model structure itself is not questioned; thus two rival models can each be successfully estimated on the same data, each taking its structure as given, but are not easily tested against each other because of familiar problems with providing a benchmark for such non-nested structures. However, the dynamic testing with which we are here concerned can naturally be used to pit rival models against each other because the fit of each to the dynamic facts can be assessed by a common metric.

There already exists a large body of work devoted to carrying out tests of this sort. It began with tests proposed by proponents of DSGE models of the Real Business Cycle type, to the effect that one should compare the correlations in the actual data with the average correlations produced by the model's dynamic simulations; the average to be computed over repeated dynamic simulations over the sample period produced by drawing repeatedly from the model's shocks. Since the focus of these models is on the business cycle, both the data and the model's shocks are detrended in some way. In this vein Christiano and Eichenbaum (1992); Burnside et al. (1993); Fève and Langot (1994) evaluate the performance of their DSGE models by using a Wald statistic checking whether or not the model's selected moments are statistically different from those in the data. Diebold et al. (1998) evaluate the performance of a model using the data's sampling variability:

they investigate whether the model's second moments lie within a prespecified confidence interval of the distribution of these moments, which is constructed by resampling the actual data. Another method uses the simulation variability from the structural model to construct the distributions of the moments of interest: the test then is whether the corresponding moments in the actual data lie within these distributions at some level of confidence (Gregory and Smith, 1991, 1993; Soderlind, 1994; Cogley and Nason, 1994). Related approaches to these compare the spectra and cross-spectra for particular groups of variables in the data and from the model; also the Impulse Response Functions (IRFs) of the data and the model for particular shocks and groups of variables.

It seems fair to say that all these comparisons suffer from a degree of arbitrariness in that moments, cross-spectra or IRFs for particular groups of variables must be chosen as the criterion of rejection while those for others are ignored. The need is for a global metric that evaluates the comparison of the model with the data across the board. In response to this need other work has developed measures of a model's overall 'distance' from the data. The first such work was that of Watson (1993). More recently Del Negro and Schorfheide (2004) have proposed deriving from the (linearised) model the VAR as restricted by the model's parameters and the data is fitted to a weighted combination of this and an unrestricted VAR. The weight on the latter measures the model's distance from the data. Different models can then be ranked according to this distance; these authors are then interested as Bayesians in combining the best features of both into an improved model based on this posterior finding of distance. One may also test the model strictly by testing the validity of the model's restrictions on the VAR.

Some researchers, most of them involved with the approaches just described (for example, DeJong et al., 2000; Schorfheide, 2000; Smets and Wouters, 2003; Fernandez-Villaverde and Rubio-Ramirez, 2004; Del Negro and Schorfheide, 2004; Corradi and Swanson, 2005) argue however that DSGE models may be heavily misspecified and can only explain certain features of the real world. Therefore, they should not be considered as legitimate objects for testing but rather evaluated only on the basis of the features they have been created to explain. However, if so they still require testing for rejection on these particular features; there will in this case be a way of isolating these features from the rest, much as 'business cycle properties' are isolated by detrending from other properties (growth etc). It is hard to see why any model should be immunised from rejection; if it makes genuinely scientific claims in the sense of Popper, then by construction it must be testable and hence rejectable[1]. Such arguments in any case would not be applicable to the model

[1]In truth it is unclear what might be meant in this context by 'misspecified'. We agree as set out in the text that there is an issue of how to allow for 'extraneous factors' in the data, such as trends, non-repeating events or measurement error; such factors are implicitly or explicitly a matter for the model's authors. Thus they may, as is usually the case, state that the model is intended only to deal with the 'business cycle' aspects of the data; if so some appropriate allowance must be made for the non-business-cycle aspects so that the model can be checked

we investigate here which is not a DSGE model but rather empirically based within a theoretical set-up derived from but not totally constrained by a DSGE model.

It is at this point that we introduce our proposal. The basic idea is to set the model, $\mathcal{M}$, up as a null hypothesis under which provisionally the model is regarded as the true description of the data, so that its structural residuals are also the true errors. Under this null, the random parts of the residuals are bootstrapped and the model simulated with these to generate a large number of sample replications, or pseudo-samples (the 'world of the model'). A descriptive time-series model, $\mathcal{T}$, typically a VAR, is then chosen for the actual data sample so that it describes the data closely and parsimoniously; such a descriptive time-series model may be consistent with many different nonlinear structural models, thus can be regarded as 'atheoretical'. The implications of $\mathcal{M}$ for the $\mathcal{T}$ that we ought to observe in the data under the null hypothesis of $\mathcal{M}$, can be discovered by estimating the same VAR on each of the pseudo-samples generated from the structural model; this will give us the distribution of the VAR's parameters according to the structural model. We can then test whether the VAR estimated on the actual data sample lies within this distribution at some level of confidence.

In the language of indirect inference, the VAR is the 'auxiliary model' to be used as the vehicle with which to estimate and test the structural model. Meenagh et al. (2008) explain how our procedure is derived from the method of indirect inference[2]. This method uses the auxiliary model

---

against the business cycle data.

Alternatively these researchers may be claiming that it is difficult to allow satisfactorily for such extraneous factors. In this case the presence of these factors unadjusted in the data would cause rejection of the model for sure; thus one is interested in whether one model is less or more distant from this unavoidably distorted data. This is a possible justification of using a distance measure where rejection is displaced by merely relative ranking. However it falls short of the requirements of science; a counsel of scientific merit would be to permit a clearcut rejection by specifying how the hypothesis is to be applied empirically (relevant features) and how not (extraneous features).

[2] The following is adapted from their explanantion. Let $x_t(\theta)$ be an $m \times 1$ vector of simulated time series dependent on the $k \times 1$ parameter vector θ and let y_t be the actual data. We assume that $x_t(\theta)$ is generated from a structural model. We assume that there exists a particular value of θ given by θ_0 such that $\{x_t(\theta_0)\}_{s=1}^{S}$ and $\{y_t\}_{t=1}^{T}$ share the same distribution, where $S = cT$ and $c \geq 1$. Thus the null hypothesis is $H_0 : \theta = \theta_0$.

Let the likelihood function defined for $\{y_t\}_{t=1}^{T}$, which is based on the auxiliary model, be $\mathcal{L}_T(y_t; a)$. The maximum likelihood estimator of a is then

$$a_T = \arg\max \mathcal{L}_T(y_t; a)$$

The corresponding likelihood function based on the simulated data $\{x_t(\theta_0)\}_{s=1}^{S}$ is $\mathcal{L}_T[x_t(\theta_0); \alpha]$. Let

$$\alpha_S = \arg\max \mathcal{L}_T[x_t(\theta_0); \alpha]$$

Define the continuous $p \times 1$ vector of functions $g(a_T)$ and $g(\alpha_S)$ and let $G_T(a_T) = \frac{1}{T}\Sigma_{t=1}^{T} g(a_T)$ and $G_S(\alpha_S) = \frac{1}{S}\Sigma_{s=1}^{S} g(\alpha_S)$. We require that $a_T \to \alpha_S$ in probability and that $G_T(a_T) \to G_S(\alpha_S)$ in probability for each θ. If $x_t(\theta)$ and y_t are stationary and ergodic then these hold *a.s.*, see Canova (2005). It then follows that on the null hypothesis, $E[g(a_T) - g(\alpha_S)] = 0$.

Thus, given an auxiliary model and a function of its parameters, we may base our test statistic for evaluating the structural model on the distribution of $g(a_T) - g(\alpha_S)$ using the Wald statistic

$$[g(a_T) - g(\alpha_S)]' W [g(a_T) - g(\alpha_S)]$$

where $W = \Sigma_g^{-1}$ and Σ_g is the covariance matrix of the (quasi) maximum likelihood estimates of $g(\alpha_S)$ which is obtained using a bootstrap simulation. The auxiliary model is a time-series model- here a VAR of varying type- and the function $g(.)$ consists of the impulse response functions of the VAR. In what follows we specialise the function $g(.)$ to (.); thus we base the test on a_T and a_S, the VAR parameters themselves.

Non-rejection of the null hypothesis is taken to indicate that the dynamic behaviour of the structural model is not significantly different from that of the actual data. Rejection is taken to imply that the structural model is incorrectly specified. Comparison of the impulse response functions of the actual and simulated data should then to describe the data- such as our time-series representation here- and estimates the parameters of the structural model of interest as those under which this model can replicate the behaviour of the auxiliary model most accurately according to a criterion of 'closeness'- see Gregory and Smith (1991, 1993), Gourieroux et al. (1993), Gourieroux and Monfort (1995) and Canova (2005). The method can also be used to evaluate the closeness of a given model; in effect this arrests the method before estimation proceeds further. This is relevant as here when we are interested in the behaviour of structural models whose structure is rather precisely specified by theory.

The paper is organized as follows. In the next section 2 we give a detailed description of the assessment process. Section 3 discusses our application to a largescale macroeconomic model of the UK. Our conclusion follows.

2 Proposed Procedure

We assume the macroeconomic model, $\mathcal{M}$, forms a system of equations, usually nonlinear: we are concerned with the whole class of structural macroeconomic models, including DSGE models and models of aggregate demand and supply derived from them in an approximative way. We represent the model quite generally as

$$y_t = f[y_t, (L)y_t, x_t, (L)x_t, y_t^e, (L)y_t^e; \theta; \varepsilon_t] \qquad t = 1, 2, ..., n \tag{1}$$

$$x_t = A(L)x_t + \in_t \tag{2}$$

where y_t is the vector of the endogenous variables, x_t is that of the exogenous variables (assumed to follow linear univariate time-series processes), (L) is the lag series $L, L^2, L^3, ..., L^k$ (where k is the maximum lag length), y_t^e is the vector of expectations (usually rational based on t-information), θ is the vector of parameters and ε_t is the vector of structural equation errors; that is to say, the errors when the endogenous variables on the right hand side are at their true values.

The proposed evaluation process is divided into three steps.

Step 1: Determination of the errors of the economic model conditional on the actual data.

In this step the structural model is solved and the errors ε_t implied by the theory conditional on the historical data are calculated; a variety of approximation methods are nowadays available for solving such models[3]. Normality of the errors is not in general assumed but it is usually assumed

reveal in what respects the structural model differs.

[3]If $\mathcal{M}$ is solved by using perturbation methods (second order approximation) then a closed form expression regarding these errors has been derived and it is available upon request. The model here is solved by using a solution method for nonlinear dynamic models with rational expectations similar to the one described by Fair and Taylor (1983); A detailed discussion can be found in Minford and Webb (2005). The method solves the model in its nonlinear form by setting future errors to their expected future values. If the error in approximating the expected

that the dimension of the vector of the stochastic processes ε_t that govern the system is the same as the dimension of the vector of the observable variables y_t. If this is not the case, as occurs with many DSGE models where the menu of shocks is restricted (for example in RBC models where usually the only error specified is the productivity shock), then the researcher is assuming that other errors are non-stochastic and are thus held constant in stochastic simulation. An example of this could be a war, which would not be expected to recur in another sample to which the model would apply. Other examples would be measurement error and trend. Then the researcher adjusts the data for the effect of such 'exogenous shocks', perhaps by adding dummy variables, or, by simply stripping out their estimated effects from the data. The remaining part of the data is assumed to represent the stochastic elements of interest; this is compared with the structural model as disturbed by the supposed stochastic shocks. Thus both the model and the data are purged of such effects, as 'extraneous factors' in the sense of our discussion above.

Step 2: The distributions of interest conditional on the structural model (the null hypothesis). According to the null hypothesis the $\{\varepsilon_t\}_{t=1}^{T}$ errors are omitted variables, modelled by autoregressive processes of identically and independently distributed (iid) shocks ε_t^* extracted as the residual from their autoregressive processes. Their empirical distribution, F_{ε^*}, is assumed to be given by the actual sample of the residuals; its variance-covariance matrix is therefore the actual one, Σ_{ε^*}. Simulations of $\mathcal{M}$ are generated from it by using standard resampling techniques for iid data. The empirical distribution of y_t conditional on the structural model $\mathcal{M}$ can then be approximated through simulation[4].

variables due to the second and higher order moments is constant, then the stochastic simulations will be accurate. Further work is needed on how this and other methods of solution compare in their approximative properties.

[4]For instance, let $\left\{\left\{\tilde{y}_t^{*,j}\right\}_{t=1}^{T}\right\}_{j=1}^{J}$ stand for the new pseudo data set. The distribution of any statistic or population moments of interest, $S_{\tilde{y}}$, conditional on $\mathcal{M}$, $f\left(S_{\tilde{y}} \mid \mathcal{M}\right)$, is then easily derived. For example we may wish to focus on the VAR properties of the data; since the conditional distribution of the actual data is in general well approximated by a VAR with iid residuals. Using the companion matrix a VAR(p), $\tilde{y}_t = A_1\tilde{y}_{t-1} + ... + A_p\tilde{y}_{t-p} + u_t$, can be written as VAR(1), $\tilde{Y}_t = A_1\tilde{Y}_{t-1} + U_t \equiv \begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{t-1} \\ \vdots \\ \tilde{y}_{t-p+1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & \cdots & A_p \\ I_{dy} & 0_{dy} & \cdots & 0_{dy} \\ \vdots & \ddots & \cdots & \vdots \\ 0_{dy} & 0_{dy} & \cdots & 0_{dy} \end{bmatrix} \begin{bmatrix} \tilde{y}_{t-1} \\ \tilde{y}_{t-2} \\ \vdots \\ \tilde{y}_{t-p} \end{bmatrix} + \begin{bmatrix} u_t \\ 0_{dy} \\ \vdots \\ 0_{dy} \end{bmatrix}$ and the second moments are given by $\Sigma_Y = A\Sigma_Y A' + \Sigma_U$ or $vec(\Sigma_Y) = \left(I_{(dyp)^2} - A\otimes A\right)^{-1} vec(\Sigma_U)$ and $\Psi_Y(i) = A^i\Sigma_Y$. Given $\left\{\left\{\tilde{y}_t^{*,j}\right\}_{t=1}^{T}\right\}_{j=1}^{J}$, $f\left(\Sigma_Y \mid \mathcal{M}\right)$ and $f\left(\Psi_Y(i) \mid \mathcal{M}\right)$ are derived by fitting at each $\left\{\tilde{y}_t^{*,j}\right\}_{t=1}^{T}$ a VAR(p) and calculating $\Sigma_Y^{*,j}$ and $\Psi_Y(i)^{*,j}$ by using the above given equations. Notice that this procedure does not require the mapping from the structural, θ, to reduced form, γ, parameters to be established analytically. $\mathcal{M}$ is nonlinear: hence for historical data $\mathcal{T}$ needs to capture these nonlinearities and thus is probably nonlinear itself. Now the only way to establish the connection between θ and γ is via simulation (this is the idea introduced by Smith, 1993). For instance, if interest or exchange rates are elements of $\tilde{y}_t$ then a linear VAR(p) is not sufficient to capture the higher order dynamics, usually, displayed by these series. This requires $\mathcal{T}$ being a VAR(p) with a time varying covariance matrix for its error vector. In this cases $\mathcal{Q}$ is derived by fitting $\mathcal{T}$ into $\left\{\tilde{y}_t^{*,j}\right\}_{t=1}^{T}$ and then $\hat{\gamma}_T^{*,j} = \arg\max_{\gamma\in\mathbf{\Gamma}} L^{\mathcal{T}}\left(\tilde{y}_t^{*,j};\gamma\right)$ is used for the calculation of $\tilde{S}_y^{*,j}$ (where $L^{\mathcal{T}}$ is the log-likelihood of $\mathcal{T}$ and, additionally, Theodoridis (2006, theorem 4.2.1) in the line drawn by Stinchcombe and White (1992) and White (1994) proves that $\gamma_T^{*,j}$ is a measurable function).

Step 3: Inference Step: a Wald statistic

In this last step we check whether $\mathcal{T}$ could have been generated by $\mathcal{M}$. We do this via a Wald statistic, to check whether the joint set of parameters $\widehat{\gamma}$ in $\mathcal{T}$ lie within the 95% (say) bands of the distribution for γ obtained from the sampling distribution of $\mathcal{M}$ (i.e. from the pseudo-sample estimates of γ).

To obtain the individual parameter confidence intervals (say at 95%) of γ, we find the two 2.5% bootstrap tails for each taken individually. However, the joint confidence interval for the whole set of γ requires the bootstrap combinations for γ to be ordered around their mean: if there are p parameters, this is a p-dimensional ordering.

To establish this ordering we compute the square of the Mahalanobis distance

$$(\hat{\gamma} - \gamma^o)'\Sigma_{\gamma}^{-1}(\hat{\gamma} - \gamma^o) \tag{3}$$

(where Σ_{γ} is the Quasi Maximum Likelihood Covariance Matrix of $\hat{\gamma}$) for the combinations, and order them in ascending distance. This then allows us to compute the 5% bootstrap 'multi-variate contour' — i.e. the parameter combinations beyond which lie the 5% of bootstraps generating the highest distance[5]. The Wald statistic, which we call the M(odel)-metric, is then the percentile value of $\widehat{\gamma}$ estimated on $\mathcal{T}$ in terms of the above bootstrap distribution; as usual a critical percentile value for it can then be chosen as the rejection boundary.

2.1 Discussion: Potential pitfalls in the process:

The process we have described contains potential pitfalls.

First, the structural model $\mathcal{M}$ which is treated as the null hypothesis to be tested can only be solved and simulated approximately; thus strictly we can only test the model in this approximated form. In Appendix A we address this issue and show that under particular assumptions we may regard our tests also as applying to the underlying structural model[6].

Second, we generate its sampling distributions via the bootstrap according to which we take the sample model errors (which under the null are the true sample errors) as representing the true distribution; under bootstrap theory we know that by taking enough bootstraps we can reduce the bootstrap representation error. However there is always some possibility of circumstances in the set-up being such that due to this the results could be misleading. There is no general way of addressing this issue since we cannot know the true distribution of the errors under the null

[5] The sensitivity of this ordering to an asymmetric distribution is an issue for further research: we are investigating the use of numerical methods directly on the bootstrap distribution to locate the 5% contour.

[6] All the appendices can be found at `http://www.cardiff.ac.uk/carbs/econ/minfordp/`.

— the sample errors are all we have. In Montecarlo simulation exercises one may check whether problems arise, were the true distribution to be of a particular sort; clearly this is an important area for future research.

Third, we choose a time-series model $\mathcal{T}$ based on the best fit to the data, subject only to the restriction that it must not violate the null hypothesis $\mathcal{M}$ in any general way that can be identified (for example in its order of integration)[7]. The idea is that this time-series representation can be used to describe the data neutrally between rival structural null hypotheses that are possible alternative candidates as models. The time-series model $\mathcal{T}$ thus takes the place of the 'dynamic facts' that a structural model must fit. However, the choice of this representation is not necessarily unique in practice as several different ones may not differ materially in fit, yet may differ in their relative capacity to reject different structural models. Again we leave this for further research.

3 An Application — testing a model of the UK 1979–2003

In this section we carry out a test of a macro model, $\mathcal{M}$, that has been used to forecast and analyze the UK economy since 1979. After a brief description of this model, we look first for a $\mathcal{T}$ that describes the historical data in a satisfactory way. Second, we review the model's performance in various dimensions, then find the Wald statistic or M-metric studied in the last section.

3.1 The model to be tested ($\mathcal{M}$)

The Liverpool Model of the UK economy is the one we examine as $\mathcal{M}$. This is an Aggregate Demand and Supply model derived by a series of approximations from the maximising conditions on consumers and firms (equations similar to these can be derived from a DSGE model in the manner of McCallum and Nelson (1999, 2000)). The first-order conditions of consumers and investors have been approximated before being used in the goods market clearing equation (an 'IS' curve); the production function is represented by a competitive price-cost equation; there are Fischer-style overlapping wage contracts that impart to wages short-lived (4-quarter) nominal rigidity; it is furthermore an open economy version of such a model. In recent work a new FIML algorithm developed in Cardiff University (Minford and Webb, 2005) has been used to reestimate the model parameters: it turns out that the new estimates are little different from the model's original ones, based partly on single-equation estimates, partly on calibration from simulation properties.

[7]The method proposed by Del Negro and Schorfheide (2004) discussed above gets around this problem by restricting $\mathcal{T}$ by $\mathcal{M}$ via a mapping $\mathcal{Q}$. This involves the linearisation of $\mathcal{M}$ after which this mapping can be derived uniquely. Then it is possible to estimate the resulting restricted $\mathcal{T}$ and test the restrictions imposed by $\mathcal{M}$. Here we have stressed the nonlinearity of the structural model as potentially causing unreliability in the linearized mapping.

The model (a full account of the original model can be found in Minford (1980), while a listing of the most recent version is in (Minford and Webb, 2005)) has been used in forecasting continuously since 1979, and is now one of only two in that category. The other is the NIESR model, which however has been frequently changed in that period: the only changes in the Liverpool Model were the introduction in the early 1980s of supply-side equations (to estimate underlying or equilibrium values of unemployment, output and the exchange rate) and the shift from annual data to a quarterly version in the mid-1980s. In an exhaustive comparative test of forecasting ability over the 1980s, Andrews et al. (1996) showed that out of three models extant in that decade — Liverpool, NIESR, and LBS — the forecasting performance of none of them could 'reject' that of the others in non-nested tests, suggesting that the Liverpool Model during this period was, though a newcomer, at least no worse than the major models of that time. For 1990s forecasts no formal test is available, but the LBS model stopped forecasts and in annual forecasting post-mortem contests NIESR came top in two years, Liverpool in three. In terms of major UK episodes, Liverpool model forecasts successfully predicted the sharp drop in inflation and the good growth recovery of the early 1980s. From the mid-1980s they rightly predicted that the underlying rate of unemployment was coming down because of supply-side reforms and that unemployment would in time fall steadily in consequence. Then they identified the weakness of UK membership of the ERM and its likely departure because of the clash between the needs of the UK economy and those of Germany leading the ERM at the time of German Reunification. After leaving the ERM they forecast that inflation would stay low and that unemployment would fall steadily from its ERM-recession peak back into line with the low underlying rate — as indeed was the case. Thus we would suggest that the Liverpool Model has a reasonable forecasting record[8].

Essentially we can summarize our interest in using this model as a test for these methods as due to the fact that (a) it has been carefully estimated on the UK data (b) it has a reasonable degree of micro-foundedness (c) it also has a fair record of fitting the data dynamically, both within sample and in forecasting. Hence we would hope that it would not be entirely rejected by the 'dynamic data'; we might therefore learn something about the practical application of the method described here.

3.1.1 The Time series representation of the data 1979–2003 ($\mathcal{T}$)

The descriptive process, $\mathcal{T}$, is a VAR set up for five key variables: output (GDP at factor cost), unemployment, inflation (Consumer Price Index), nominal interest rates (Nominal deposit interest rate with local authorities, 3 month) and nominal exchange rates (Trade-weighted exchange rate)

[8]The model's equations are listed at appendix A.2.

(see Minford et al., 2004, for a detailed description of the model and the data used here). We chose to detrend the data by differencing; other methods of detrending impose an arbitrary degree of smoothing on the data to find the 'trend' and it is not clear what exactly they remove from the data — thus the 'cyclical' elements may have been massaged in an unknown way. Differencing removes both deterministic and stochastic trend; the history of the variable is lost however, so this method too is far from ideal. Further work on the appropriate way to treat trend is both necessary and also underway, but lies beyond the scope of this paper.

The policy regime followed during the sample period was not constant. It began in 1979 with a money supply targeting regime. It switched in 1986 to an exchange rate targeting regime. Then from 1992 it switched finally to an inflation targeting regime. We experimented with two main representations of this regime change in this study, each allowing the parameter vector γ to vary across regimes in a different way. In the first γ follows a Markov process where there are three regimes with state-dependent switches between them. In the second γ changes with each regime, switching deterministically in 1986 and 1992. The length of $y_t \in \mathbb{R}^5$ (where $\mathbb{R}$ denotes the real line) is large and very quickly the number of the estimated parameters becomes huge while, on the other hand, the macroeconomic data is limited. Due to this data limitation the Information Matrix Equality Test (see White, 1994, Chapter 11) cannot be applied and the choice of the most plausible $\mathcal{T}$ is founded on the basis of some diagnostic tests on the estimated residuals, selection criteria and computational simplicity.

3.1.2 Model $\mathcal{T}_1$: VAR(1) with Markov-Switching between 3 Regimes:

The first time series model estimated is a VAR(1) model, $\Delta y_t = c + A_1 \Delta y_{t-1} + u_t$, where the variance-covariance matrix of the errors processes is regime, κ_t, dependent, $u_t \sim N(0, \sum_{i=1}^{3} \Sigma_{u,s}\xi_t)$, and this regime follows a first order Markov Process, $P\{\kappa_t = j|\kappa_{t-1} = i, \kappa_{t-1} = k, ...\} = P\{\kappa_t = j|\kappa_{t-1} = i\} = p_{ij}$. ξ_t is a Markov Chain

$$\xi_t = \begin{cases} (1,0,0)' & \text{when} \quad s_t = 1 \\ (0,1,0)' & \text{when} \quad s_t = 2 \\ (0,0,1)' & \text{when} \quad s_t = 3 \end{cases}$$

which follows a VAR(1) process

$$\xi_{t+1} = P\xi_t + w_{t+1}$$

where $P = \begin{bmatrix} p_{11} & p_{21} & p_{31} \\ p_{12} & p_{22} & p_{32} \\ p_{13} & p_{23} & p_{33} \end{bmatrix}$ is the matrix of the transition probabilities. This model, $\mathcal{T}_1$, also identified as MSH(3)–VAR(1) is estimated using techniques described by Krolzig (1997).

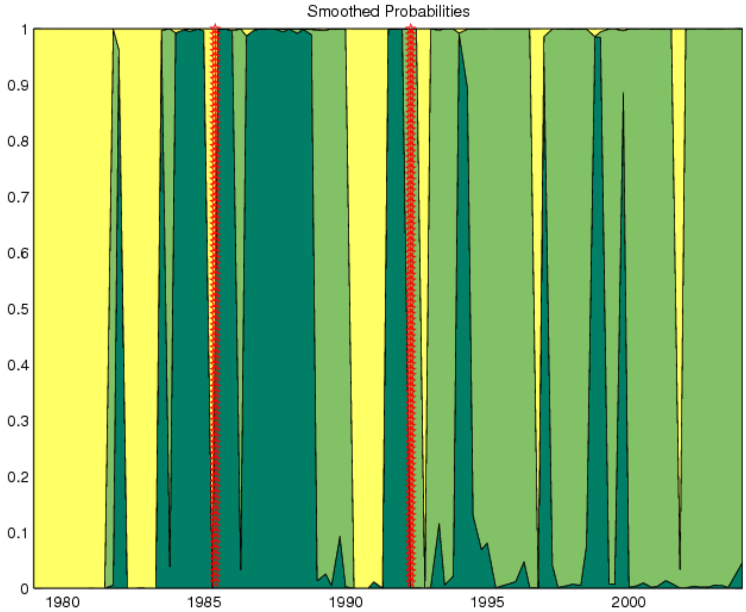

Figure 1: **Pattern of Smoothed Probabilities for three regimes** (yellow: Reg. 1; dark green: Reg. 2; light green: Reg. 3)

It seems to estimate the policy regime turning points fairly accurately: thus the pattern of the smoothed probabilities (Figure 1) broadly picks up these monetary policy switches. For instance, it is clear from the diagram that the area after 1992 (second, from the LHS, vertical red line) is dominated by the existence of one state (light green area), which could be viewed as the inflation targeting regime introduced after that date. The monetary targeting regime (yellow area) similarly dominates the period before 1985Q4 (as indicated by the first vertical red line) but ends a few quarters earlier than it is expected to. Table 1 shows some of the properties of the estimated residuals; they suggest that $\mathcal{T}_1$ is mis-specified. Nevertheless we return to it below, after considering the second time-series representation.

Table 1: **Diagnostic Tests of $\mathcal{T}_1$**

Statistics	*Chi-Square*	*DF*	*P-values*
Conditional Heteroskedasticity	1101.1	900	0
Heteroskedasticity	266.5	150	0
Q- Statistics lags (12)	385.6	275	0
Adj-Q- Statistics lags (12)	410.95	275	0

3.1.3 Model $\mathcal{T}_2$: Fixed Regime Switching:

The second time series model, $\mathcal{T}_2$, estimated here is a VAR(1) whose autoregressive and variance-covariance parameters vary across regimes in a deterministic way.

$$\Delta y_t = \begin{cases} c_1 + A_{1,1}\Delta y_{t-1} + \epsilon_t & E(\epsilon_t \epsilon_t') = \Sigma_1 & t \leq T_1 = 1986 \\ c_2 + A_{1,2}\Delta y_{t-1} + \epsilon_t & E(\epsilon_t \epsilon_t') = \Sigma_2 & T_1 < t \leq T_2 = 1992 \\ c_3 + A_{1,3}\Delta y_{t-1} + \epsilon_t & E(\epsilon_t \epsilon_t') = \Sigma_3 & t > T_3 = 1992 \end{cases}$$

This model is called an Intervention model; a detailed discussion of these models can be found in Lutkepohl (1993, Section 12.4). This representation implies that a particular stationary DGP is in operation up to a certain period, with another process generating the data after that period; here there are two such breaks and three distinct regimes. One of the advantages of this data representation process is its simplicity both from an intuitive and computational point of view. However, the estimation of this model poses some theoretical difficulties, such as consistency[9], which cannot be overcome easily. Another important characteristic of the Intervention model used here is that after an intervention the moments of the process do not reach a fixed new level immediately but only asymptotically. For instance, let us assume that the intervention appears at $t = 4$ solving backwards up to $t = 3$ we get $\Delta y_3 = \sum_{i=0}^{2} A_{1,1}^i c_1 + \sum_{i=0}^{2} A_{1,1}^i \epsilon_{t-i} + A_{1,1}^3 \Delta y_0$, for $t = 4$, $\Delta y_4 = c_2 + A_{1,2}\Delta y_3 + \epsilon_4$ and, for $t > 4$, $\Delta yt = \sum_{i=0}^{t-1} A_{1,2}^i c_2 + \sum_{i=0}^{t-1} A_{1,2}^i \epsilon_{t-i} + A_{1,2}^t \left(\sum_{i=0}^{2} A_{1,1}^i c_1 + \sum_{i=0}^{2} A_{1,1}^i \epsilon_{t-i} + A_{1,1}^3 \Delta y_0\right)$. The last term of the latter expression vanishes as $t \to \infty$ due to the stationarity assumption (with the maximum eigenvalue of $A_{1,2}$ less than one in absolute terms). This assumption, $t \to \infty$, may make sense if there is only one intervention but this is not the case here. However, this behaviour may be quite plausible in practice because a system may react slowly to an intervention (see Lutkepohl, 1993, Section 12.4.1, page:409).

The Table 2 illustrates the properties of $\mathcal{T}_2$'s standardized residuals, which do not seem to contain any structure. Additionally, Table 3, which presents the values of various selection (Akaike,

[9]The estimates of the middle regimes are functions of their sample sizes which do not tend to infinity. A convenient but not very plausible assumption requires T_i, for $i = 1, 2$, being fractions of T and they tend to infinity as $T \to \infty$.

Table 2: **Diagnostic Tests of $\mathcal{T}_2$**

Statistics	*Chi-Square*	*DF*	*P-values*
Conditional Heteroskedasticity	203.45	225	0.846
Heteroskedasticity	162.75	150	0.225
Q- Statistics lags (18)	383.1	425	0.928
Adj-Q- Statistics lags (18)	469.3	425	0.068

Table 3: **Selection Criteria**

Information Criteria	$\mathcal{T}_1$	$\mathcal{T}_2$
AIC	-31.657	-37.888
SIC	-28.140	-34.349
HQ	-30.233	-36.456

Schwartz and Hannan-Quinn) criteria, indicates that $\mathcal{T}_2$ has a decidedly better fit than $\mathcal{T}_1$. In what follows we first look in some detail at the comparison between $\mathcal{T}_2$ as the best-fitting time-series model, and $\mathcal{M}$; we then look more briefly at that between $\mathcal{T}_1$ and $\mathcal{M}$.

3.2 Comparing $\mathcal{T}_2$ and $\mathcal{M}$.

We begin by discussing the behaviour of the three-regime variables, the time-series model errors and the comparison of the second moments of $\mathcal{T}_2$ and $\mathcal{M}$. The maximum eigenvalue of $\left\{\hat{A}_{1,i}\right\}_{i=1}^{3}$ and the trace of the Variance-Covariance matrices of ϵ_t, $\left\{tr(\Sigma_{\epsilon}^{i})\right\}_{i=1}^{3}$, are calculated (Table 4). The maximum eigenvalue and the trace indicate the persistence and the uncertainty, respectively,

Table 4: **Regime Features ($\mathcal{T}_2$)**

Features/Regimes	*Monetary*	*Exchange Rate*	*Inflation*
Maximum Eigenvalue	0.451	0.933	0.68
Trace of the Covariance Matrix	0.0043	0.0012	0.0007

that the whole vector, $\Delta y_t = (\Delta \mathbf{y}_t', \Delta u_t', \Delta \pi_t', \Delta r_t', \Delta e_t')'$, displays over policy regimes. Two characteristics are worth mentioning from this table, first, the remarkable persistence that the vector of 5 variables Δy_t shows during the middle regime and, second, that the size of uncertainty falls steadily.

Turning to the moments, we use the simulation variability of $\mathcal{M}$ (as the null hypothesis) to ask whether the data description, $\mathcal{T}_2$, can reject $\mathcal{M}$. Given $\mathcal{M}$ and the actual data, y_t, the simulated distribution of $\tilde{S}_y$[10] conditional on the actual data, $f\left(\tilde{S}_y|y_t,\mathcal{M}\right)$, is readily derived and we ask whether or not S_y[11] lies within 95% of the model-simulated distribution. Column 2 of the following Tables show the moments estimated by $\mathcal{T}_2$ on the actual data. Columns 3 and 4 show the estimates of these moments derived from estimating $\mathcal{T}_2$ on the pseudo-samples bootstrapped from $\mathcal{M}$.

Table 5: **Monetary Targeting/Simulation Variability ($\mathcal{T}_2$)**

Second Moments	$\mathcal{T}_2$*'s Moments*	$\mathcal{M}$*'s Lower Bound*	$\mathcal{M}$*'s Upper Bound*	*Bootstrapped Prob-Values*
$\sigma_{\Delta y \Delta y}$	0.0089	0.0009	0.0321	0.252
$\sigma_{\Delta u \Delta y}$	-0.0009	-0.0061	0.0070	0.468
$\sigma_{\Delta \pi \Delta y}$	0.0134	-0.0078	0.0224	0.059
$\sigma_{\Delta r \Delta y}$	-0.0002	-0.0347	0.0079	0.423
$\sigma_{\Delta e \Delta y}$	-0.0006	-0.0091	0.0082	0.658
$\sigma_{\Delta u \Delta u}$	0.0026	0.0027	0.0210	0.978*
$\sigma_{\Delta \pi \Delta u}$	-0.0014	-0.0071	0.0082	0.817
$\sigma_{\Delta r \Delta u}$	-0.0003	-0.0049	0.0109	0.767
$\sigma_{\Delta e \Delta u}$	0.0004	-0.0023	0.0064	0.510
$\sigma_{\Delta \pi \Delta \pi}$	0.0327	0.0006	0.0462	0.045
$\sigma_{\Delta r \Delta \pi}$	-0.0012	-0.0336	0.0062	0.557
$\sigma_{\Delta e \Delta \pi}$	-0.0013	-0.0097	0.0081	0.782
$\sigma_{\Delta r \Delta r}$	0.0008	0.0008	0.0528	0.983*
$\sigma_{\Delta e \Delta r}$	0.0001	-0.0053	0.0132	0.535
$\sigma_{\Delta e \Delta e}$	0.0015	0.0016	0.0163	0.985*

Thus here we are directly testing the model and seeing whether the moment estimates from the data individually lie within the 95% confidence intervals implied by the model. From Tables 5 and 6 we can see that the variances of Δr_t and Δe_t cause a problem in both the first two regimes (money and exchange rate targeting) the Liverpool Model produces excessive variance of both interest rates and exchange rates; in the money targeting it also produces excessive unemployment variance. In the inflation-targeting regime it produces generally excessive variance; not merely of inflation, interest rates and the exchange rate but also of output (Table 7); thus it fails to capture the 'great moderation' in the UK whereby under inflation targeting volatility of the economy declined sharply on all fronts.

We now consider the estimates of the parameter vector, γ in $\mathcal{T}_2$ and whether individually they lie within the 95% bootstrap distributions produced by $\mathcal{M}$. These parameters represent the partial regressors of each variable on the lagged values of itself and the others. One rather striking feature of the comparison is how many of these regressor estimates individually reject $\mathcal{M}$

The failure of the structural model to match this data representation shows up mainly in the variances in all three regimes, much as we have already seen in the cruder comparison of moments.

[10] $\tilde{S}_y$ denotes the second moments of Δy_t produced by fitting $\mathcal{T}_2$ to the data simulated by $\mathcal{M}$.
[11] S_y denotes again the second moments of Δy_t produced by fitting $\mathcal{T}_2$ to the actual data now.

Table 6: **Exchange Rate Targeting/Simulation Variability ($\mathcal{T}_2$)**

Second Moments	*$\mathcal{T}_2$'s Moments*	*$\mathcal{M}$'s Lower Bound*	*$\mathcal{M}$'s Upper Bound*	*Bootstrapped Prob-Values*
$\sigma_{\Delta y \Delta y}$	0.0026	0.0008	0.1451	0.161
$\sigma_{\Delta u \Delta y}$	-0.0013	-0.0044	0.0478	0.166
$\sigma_{\Delta \pi \Delta y}$	0.0005	-0.0299	0.0284	0.584
$\sigma_{\Delta r \Delta y}$	-0.0001	-0.0081	0.0067	0.490
$\sigma_{\Delta e \Delta y}$	-0.0001	-0.3487	0.2117	0.447
$\sigma_{\Delta u \Delta u}$	0.0024	0.0019	0.0350	0.069
$\sigma_{\Delta \pi \Delta u}$	-0.0026	-0.0076	0.0099	0.140
$\sigma_{\Delta r \Delta u}$	0.0008	-0.0032	0.0027	0.750
$\sigma_{\Delta e \Delta u}$	0.0006	-0.1763	0.1823	0.594
$\sigma_{\Delta \pi \Delta \pi}$	0.0048	0.0009	0.0608	0.510
$\sigma_{\Delta r \Delta \pi}$	-0.0017	-0.0037	0.0033	0.084
$\sigma_{\Delta e \Delta \pi}$	-0.0008	-0.1380	0.1173	0.555
$\sigma_{\Delta r \Delta r}$	0.0006	0.0031	0.0200	0.000**
$\sigma_{\Delta e \Delta r}$	0.0002	-0.0253	0.0436	0.510
$\sigma_{\Delta e \Delta e}$	0.0009	0.0015	35.2037	0.012**

The model produces excessive variances, particularly of inflation and interest rates. In the first two regimes the VAR parameters in the lag matrix are generally within the 95% limits; three are outside in the first (Table 8), five in the second (Table 9). However in the third (Table 10), inflation targeting, no less than nine are outside.

When we turn to the M-metric, we find that the model is rejected at the 95% level by the data as represented by $\mathcal{T}_2$. This is shown in Table 11: the M-metric is 98.8%, well outside the critical range. Thus the model is clearly rejected. The reason is not difficult to see from inspecting the failure at the individual coefficient level: the model generates excessive variances and even if it fails less in picking up lagged responses, the percentage of these coefficients outside the 95% range is still unacceptably high.

3.3 The Impulse Responses in the data and the model

It is instructive to compare the impulse responses (to particular shocks identified in the model) that are implied by the data representation $\mathcal{T}_2$ with the 95% bounds placed on it by the structural model. We compare the impulse responses from the VAR ($\mathcal{T}_2$) when the model's restrictions are used to identify the relevant shocks to the VAR; of course the VAR itself cannot tell us what its shocks represent as in general they are linear combinations of the underlying shocks in the model. Identification of the underlying shock is done by establishing a mapping from the model shock involved to the model-implied shocks for the variables in the VAR; these shocks are then input into the VAR and the resulting movement in those variables plotted; the 95% confidence bounds

Table 7: **Inflation Rate Targeting/Simulation Variability ($\mathcal{T}_2$)**

Second Moments	*$\mathcal{T}_2$'s Moments*	*$\mathcal{M}$'s Lower Bound*	*$\mathcal{M}$'s Upper Bound*	*Bootstrapped Prob-Values*
$\sigma_{\Delta y \Delta y}$	0.0012	0.0013	0.0778	0.017**
$\sigma_{\Delta u \Delta y}$	-0.0002	-0.0047	0.0151	0.542
$\sigma_{\Delta \pi \Delta y}$	0.0001	-0.0112	0.0064	0.658
$\sigma_{\Delta r \Delta y}$	-0.0008	-0.0156	0.0183	0.443
$\sigma_{\Delta e \Delta y}$	0.0001	-0.0239	0.0257	0.443
$\sigma_{\Delta u \Delta u}$	0.0011	0.0027	0.0215	0.000**
$\sigma_{\Delta \pi \Delta u}$	-0.0004	-0.0031	0.0018	0.530
$\sigma_{\Delta r \Delta u}$	0.0031	-0.0047	0.0036	0.955
$\sigma_{\Delta e \Delta u}$	-0.0004	-0.0041	0.0071	0.191
$\sigma_{\Delta \pi \Delta \pi}$	0.0002	0.0030	0.0108	0.000**
$\sigma_{\Delta r \Delta \pi}$	-0.0011	-0.0037	0.0127	0.228
$\sigma_{\Delta e \Delta \pi}$	0.0001	-0.0113	-0.0001	0.975
$\sigma_{\Delta r \Delta r}$	0.0101	0.0014	0.0525	0.408
$\sigma_{\Delta e \Delta r}$	-0.0009	-0.0282	0.0094	0.757
$\sigma_{\Delta e \Delta e}$	0.0006	0.0076	0.0384	0.000**

for this are found from the VAR bootstraps.

In what follows we show the impulse responses generated by a supply shock (a rise in the employers' tax rate on workers) and by a nominal shock (this is a fall in the money supply under the Money targeting regime; a fall in foreign prices under Exchange rate targeting; and a temporary rise in interest rates under the Inflation targeting regime). The response of the VAR when the VAR shocks are identified via the effect of the structural model (so that in effect their first period effect is the same as that of $\mathcal{M}$) is shown by the solid line; and the 95% confidence interval for the VAR effect coming from the model bootstraps is shown by the two dotted lines. If the model is correct, we should observe that the VAR effects lie within the 95% interval. (The $\mathcal{M}$ response itself is not strictly relevant and not therefore shown). The responses are from a VAR in the differences of the variables, thus they show the effects on the rate of change of these variables to a temporary rise in the rate of change of an exogenous shock; thus they can equivalently be read as the responses of the levels of these variables to a temporary rise in the level of the shock (in effect dividing both sides by the difference operator).

As we would expect from the fact that the model is rejected by the data at the 95% level, the blue lines lie in several cases on or outside the 95% boundaries. This is particularly marked for the nominal shock in the inflation targeting regime, where all except the interest rate response lie outside. In effect the impulse response functions merely reorganise the information in the VAR lag parameters.

One interesting feature of the VAR, when the shocks are identified by this structural model, is

Table 8: **Distribution of the Parameters of $\mathcal{T}_2$ (Monetary Targeting Regime)**

Parameters	*Actual*	*Lower Bound*	*Upper Bound*	*Bootstrapped Prov. Values*	*State*
$A_{\Delta y}^{\Delta y}$	0.0128	-0.5577	0.4447	0.366	IN
$A_{\Delta y}^{\Delta u}$	-1.6212	-1.6854	1.0895	0.973	IN
$A_{\Delta y}^{\Delta \pi}$	0.1404	-0.3343	0.4907	0.356	IN
$A_{\Delta y}^{\Delta r}$	0.0385	-0.6769	0.7618	0.525	IN
$A_{\Delta y}^{\Delta e}$	0.007	-0.8879	1.0028	0.473	IN
$A_{\Delta u}^{\Delta y}$	-0.0455	-0.0839	0.226	0.923	IN
$A_{\Delta u}^{\Delta u}$	0.1606	-0.1373	0.7542	0.844	IN
$A_{\Delta u}^{\Delta \pi}$	0.0194	-0.1637	0.0577	0.116	IN
$A_{\Delta u}^{\Delta r}$	0.0165	-0.2276	0.2384	0.475	IN
$A_{\Delta u}^{\Delta e}$	-0.0033	-0.2867	0.2748	0.463	IN
$A_{\Delta \pi}^{\Delta y}$	0.1409	-0.5999	0.5317	0.267	IN
$A_{\Delta \pi}^{\Delta u}$	-2.9414	-1.4825	1.4928	0.000	OUT
$A_{\Delta \pi}^{\Delta \pi}$	-0.3617	-0.6956	0.1234	0.661	IN
$A_{\Delta \pi}^{\Delta r}$	0.6422	-0.7575	0.8289	0.045	IN
$A_{\Delta \pi}^{\Delta e}$	-1.1134	-1.1167	0.7214	0.975	OUT
$A_{\Delta r}^{\Delta y}$	-0.0536	-0.702	0.2467	0.287	IN
$A_{\Delta r}^{\Delta u}$	-0.0684	-0.8594	1.5914	0.802	IN
$A_{\Delta r}^{\Delta \pi}$	0.0666	-0.5257	0.1092	0.035	IN
$A_{\Delta r}^{\Delta r}$	0.0882	-0.8109	0.6828	0.451	IN
$A_{\Delta r}^{\Delta e}$	0.8146	-0.8272	0.5971	0.007	OUT
$A_{\Delta e}^{\Delta y}$	-0.1094	-0.3299	0.187	0.649	IN
$A_{\Delta e}^{\Delta u}$	0.0666	-0.6253	0.8872	0.676	IN
$A_{\Delta e}^{\Delta \pi}$	0.0257	-0.0461	0.3785	0.896	IN
$A_{\Delta e}^{\Delta r}$	-0.0581	-0.1803	0.4167	0.879	IN
$A_{\Delta e}^{\Delta e}$	0.2576	-0.4259	0.3434	0.050	IN
$\sigma_{\Delta y \Delta y}$	0.0002	0.0003	0.0012	0.000	OUT
$\sigma_{\Delta u \Delta y}$	-0.0001	-0.0029	-0.0005	0.000	OUT
$\sigma_{\Delta \pi \Delta y}$	0.0000	-0.0005	0.0000	0.035	IN
$\sigma_{\Delta r \Delta y}$	0.0000	-0.0004	0.0004	0.540	IN
$\sigma_{\Delta e \Delta y}$	0.0000	-0.0003	0.0010	0.809	IN
$\sigma_{\Delta u \Delta u}$	0.0025	0.0023	0.0112	0.973	IN
$\sigma_{\Delta \pi \Delta u}$	-0.0002	-0.0001	0.0013	0.980	OUT
$\sigma_{\Delta r \Delta u}$	-0.0002	-0.0011	0.0012	0.772	IN
$\sigma_{\Delta e \Delta u}$	0.0005	-0.0025	0.0014	0.153	IN
$\sigma_{\Delta \pi \Delta \pi}$	0.0000	0.0002	0.0009	0.000	OUT
$\sigma_{\Delta r \Delta \pi}$	0.0000	0.0000	0.0009	0.963	IN
$\sigma_{\Delta e \Delta \pi}$	0.0000	-0.0011	0.0001	0.059	IN
$\sigma_{\Delta r \Delta r}$	0.0002	0.0004	0.0032	0.000	OUT
$\sigma_{\Delta e \Delta r}$	-0.0001	-0.0012	0.0004	0.428	IN
$\sigma_{\Delta e \Delta e}$	0.0016	0.0011	0.0046	0.844	IN

Table 9: **Distribution of the Parameters of $\mathcal{T}_2$ (Exchange Rate Targeting Regime)**

Parameters	*Actual*	*Lower Bound*	*Upper Bound*	*Bootstrapped Prov. Values*	*State*
$A_{\Delta y}^{\Delta y}$	0.2194	-0.4603	0.2233	0.030	IN
$A_{\Delta y}^{\Delta u}$	-1.1006	-1.0583	1.0384	0.985	OUT
$A_{\Delta y}^{\Delta \pi}$	0.0888	-0.4245	0.3437	0.312	IN
$A_{\Delta y}^{\Delta r}$	-0.0236	-1.2564	1.1196	0.515	IN
$A_{\Delta y}^{\Delta e}$	0.1560	-0.5401	0.6296	0.337	IN
$A_{\Delta u}^{\Delta y}$	-0.0275	-0.0839	0.1553	0.728	IN
$A_{\Delta u}^{\Delta u}$	0.5950	0.1412	0.8215	0.500	IN
$A_{\Delta u}^{\Delta \pi}$	0.0086	-0.1607	0.0376	0.079	IN
$A_{\Delta u}^{\Delta r}$	-0.0401	-0.3664	0.3115	0.629	IN
$A_{\Delta u}^{\Delta e}$	-0.0087	-0.1498	0.2690	0.748	IN
$A_{\Delta \pi}^{\Delta y}$	0.2565	-0.3542	0.1614	0.003	OUT
$A_{\Delta \pi}^{\Delta u}$	-1.3935	-0.4574	0.9537	0.000	OUT
$A_{\Delta \pi}^{\Delta \pi}$	-0.4377	-0.7264	-0.1493	0.525	IN
$A_{\Delta \pi}^{\Delta r}$	0.2642	-1.1670	1.1637	0.300	IN
$A_{\Delta \pi}^{\Delta e}$	-0.9034	-0.3492	0.3443	0.000	OUT
$A_{\Delta r}^{\Delta y}$	-0.0715	-0.1031	0.0769	0.936	IN
$A_{\Delta r}^{\Delta u}$	0.0218	-0.2266	0.2782	0.460	IN
$A_{\Delta r}^{\Delta \pi}$	0.0619	-0.1540	0.0900	0.072	IN
$A_{\Delta r}^{\Delta r}$	0.0817	-0.4519	0.4811	0.300	IN
$A_{\Delta r}^{\Delta e}$	0.6310	-0.1487	0.1007	0.000	OUT
$A_{\Delta e}^{\Delta y}$	-0.0842	-0.3084	0.1680	0.443	IN
$A_{\Delta e}^{\Delta u}$	-0.0815	-0.6240	0.9357	0.884	IN
$A_{\Delta e}^{\Delta \pi}$	0.0272	-0.0163	0.3365	0.928	IN
$A_{\Delta e}^{\Delta r}$	0.0095	-0.2134	0.4592	0.723	IN
$A_{\Delta e}^{\Delta e}$	0.2444	-0.2999	0.4227	0.173	IN
$\sigma_{\Delta y \Delta y}$	0.0002	0.0007	0.0020	0.000	OUT
$\sigma_{\Delta u \Delta y}$	-0.0001	-0.0047	-0.0014	0.000	OUT
$\sigma_{\Delta \pi \Delta y}$	0.0000	-0.0007	0.0002	0.168	IN
$\sigma_{\Delta r \Delta y}$	0.0000	-0.0018	0.0012	0.478	IN
$\sigma_{\Delta e \Delta y}$	-0.0001	-0.0005	0.0010	0.715	IN
$\sigma_{\Delta u \Delta u}$	0.0033	0.0048	0.0161	0.000	OUT
$\sigma_{\Delta \pi \Delta u}$	-0.0002	-0.0005	0.0018	0.903	IN
$\sigma_{\Delta r \Delta u}$	-0.0003	-0.0037	0.0047	0.658	IN
$\sigma_{\Delta e \Delta u}$	0.0005	-0.0025	0.0020	0.290	IN
$\sigma_{\Delta \pi \Delta \pi}$	0.0000	0.0009	0.0023	0.000	OUT
$\sigma_{\Delta r \Delta \pi}$	0.0000	-0.0036	0.0025	0.342	IN
$\sigma_{\Delta e \Delta \pi}$	0.0000	-0.0015	-0.0002	0.005	OUT
$\sigma_{\Delta r \Delta r}$	0.0003	0.0072	0.0544	0.000	OUT
$\sigma_{\Delta e \Delta r}$	-0.0002	-0.0026	0.0037	0.748	IN
$\sigma_{\Delta e \Delta e}$	0.0029	0.0017	0.0054	0.458	IN

Table 10: **Distribution of the Parameters of $\mathcal{T}_2$ (Inflation Targeting Regime)**

Parameters	*Actual*	*Lower Bound*	*Upper Bound*	*Bootstrapped Prov. Values*	*State*
$A^{\Delta y}_{\Delta y}$	0.3733	-0.3820	0.1778	0.007	OUT
$A^{\Delta u}_{\Delta y}$	-0.7354	-0.6356	0.7531	0.985	OUT
$A^{\Delta \pi}_{\Delta y}$	0.0561	-0.5523	0.4018	0.295	IN
$A^{\Delta r}_{\Delta y}$	-0.0125	-0.5889	0.5776	0.559	IN
$A^{\Delta e}_{\Delta y}$	-0.0093	-0.4048	0.6115	0.708	IN
$A^{\Delta y}_{\Delta u}$	-0.0088	-0.0794	0.1243	0.564	IN
$A^{\Delta u}_{\Delta u}$	0.6629	0.3482	0.8514	0.488	IN
$A^{\Delta \pi}_{\Delta u}$	0.0058	-0.1529	0.0786	0.230	IN
$A^{\Delta r}_{\Delta u}$	-0.0399	-0.1886	0.1735	0.703	IN
$A^{\Delta e}_{\Delta u}$	-0.0689	-0.1449	0.2478	0.901	IN
$A^{\Delta y}_{\Delta \pi}$	0.2715	-0.1531	0.0392	0.000	OUT
$A^{\Delta u}_{\Delta \pi}$	-1.0064	-0.0661	0.3944	0.000	OUT
$A^{\Delta \pi}_{\Delta \pi}$	-0.4491	-0.7150	-0.2804	0.300	IN
$A^{\Delta r}_{\Delta \pi}$	0.2122	-0.1375	0.1683	0.012	OUT
$A^{\Delta e}_{\Delta \pi}$	-1.1928	-0.0594	0.4041	0.000	OUT
$A^{\Delta y}_{\Delta r}$	-0.0099	-0.1433	0.0528	0.470	IN
$A^{\Delta u}_{\Delta r}$	0.0888	-0.1325	0.2723	0.131	IN
$A^{\Delta \pi}_{\Delta r}$	0.0370	-0.2078	0.0899	0.094	IN
$A^{\Delta r}_{\Delta r}$	0.1174	-0.2917	0.2477	0.097	IN
$A^{\Delta e}_{\Delta r}$	0.3020	-0.1228	0.2146	0.000	OUT
$A^{\Delta y}_{\Delta e}$	-0.0501	-0.2451	0.0995	0.267	IN
$A^{\Delta u}_{\Delta e}$	-0.1224	-0.3002	0.5749	0.956	IN
$A^{\Delta \pi}_{\Delta e}$	0.0128	0.2657	0.8698	0.000	OUT
$A^{\Delta r}_{\Delta e}$	0.0212	-0.1720	0.2328	0.542	IN
$A^{\Delta e}_{\Delta e}$	0.2147	-0.4580	0.1212	0.010	OUT
$\sigma_{\Delta y \Delta y}$	0.0002	0.0009	0.0028	0.000	OUT
$\sigma_{\Delta u \Delta y}$	-0.0001	-0.0060	-0.0020	0.000	OUT
$\sigma_{\Delta \pi \Delta y}$	0.0000	-0.0011	0.0006	0.364	IN
$\sigma_{\Delta r \Delta y}$	0.0000	-0.0010	0.0009	0.550	IN
$\sigma_{\Delta e \Delta y}$	0.0000	-0.0005	0.0013	0.696	IN
$\sigma_{\Delta u \Delta u}$	0.0023	0.0062	0.0154	0.000	OUT
$\sigma_{\Delta \pi \Delta u}$	-0.0001	-0.0018	0.0014	0.408	IN
$\sigma_{\Delta r \Delta u}$	-0.0002	-0.0026	0.0030	0.644	IN
$\sigma_{\Delta e \Delta u}$	0.0002	-0.0018	0.0021	0.579	IN
$\sigma_{\Delta \pi \Delta \pi}$	0.0000	0.0033	0.0076	0.000	OUT
$\sigma_{\Delta r \Delta \pi}$	0.0000	-0.0022	0.0018	0.364	IN
$\sigma_{\Delta e \Delta \pi}$	0.0000	-0.0062	-0.0022	0.000	OUT
$\sigma_{\Delta r \Delta r}$	0.0002	0.0047	0.0333	0.000	OUT
$\sigma_{\Delta e \Delta r}$	-0.0001	-0.0016	0.0026	0.767	IN
$\sigma_{\Delta e \Delta e}$	0.0021	0.0038	0.0091	0.000	OUT

Table 11: **M-metric (Wald statistic) for Model using $\mathcal{T}_2$**

CI	*Simulated Statistic*	*Actual Statistic*
95%	1.21	
96%	1.325	
97%	1.574	
98%	1.699	
99%	***1.965***	***1.9598 (98.76%)***
100%	5.918	

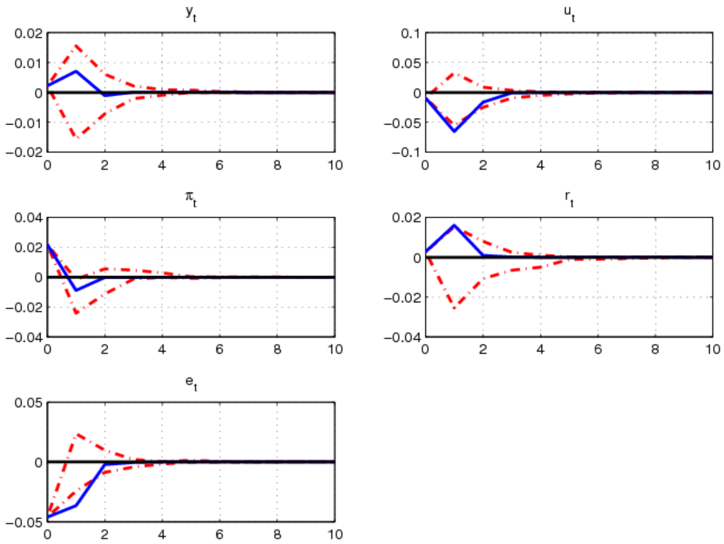

Figure 2: **Monetary Targeting Regime / Nominal Shock**

that it does not imply 'hump-shaped' responses of output and inflation to nominal shocks — as it has been widely suggested that VAR data imply (eg, Christiano et al., 2005). Such responses were found by Christiano et al. on US data when they identified the shocks using their particular model in which interest rates are assumed to have no contemporaneous (quarterly) effect on output or inflation; however in the model here that is not assumed. What this implies is that VARs cannot unambiguously determine whether there are or are not particular shapes of response; this can only be determined under the identification scheme supplied by the model. Thus for example it is simply not the case that hump shapes 'are implied' by the VAR.

3.4 Comparing $\mathcal{T}_1$ and $\mathcal{M}$.

In the VAR data representation $\mathcal{T}_1$ we assumed that the model's variances and covariances switched in a Markov manner between regimes while the VAR lag parameters remained unchanged. This representation did not fit the data as well as $\mathcal{T}_2$ which we have just examined where we divided

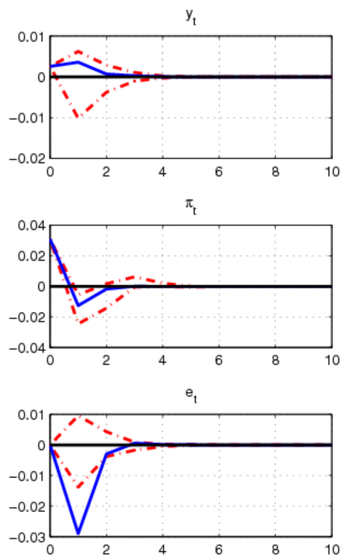

Figure 3: **Exchange Rate Targeting Regime / Supply Shock**

the data into three separate VARs. However it is of interest in representing the regimes as less clear-cut; in effect it allows for short bursts of 'other regime' behaviour within each regime period, as if uncertainty prevailed over whether a given regime in progress was sustainable. The structural model necessarily is solved under the assumption of fixed regime switches; but these regimes may not always have been credible at all times. $\mathcal{T}_1$ identifies regimes by the size of the variances associated with them; thus regime 1 (monetary targeting) has the highest variances, regime 3 (inflation targeting) the lowest, and regime 2 (exchange rate targeting) lies between them.

Predominantly these $\mathcal{T}_1$-regimes correspond to the actual ones, but with punctuation, as can be seen in Figure 1. Thus $\mathcal{T}_1$ resembles a GARCH framework but where variances shift according to the state of the economy and not past variances.

If we take $\mathcal{T}_1$ as the representation we find that the structural model is less at variance with the dynamic facts. Tables 12 and 13 show the lag and the covariance parameters respectively. There is a marked reduction in rejected parameters, relative to the rejections for $\mathcal{T}_2$. In particular only one covariance parameter is rejected, that for inflation variance in regime 2; the model produces excessive variance but only modestly so. Turning to the M-metric in Table 14, we find that it is just below 95%, and so as we might have expected from the few individual parameter rejections, the model as a whole is just accepted.

This comparison of the structural model with $\mathcal{T}_1$ shows that if the restrictions in the model are relaxed to allow more regime uncertainty then the model is not rejected by the dynamic data. However the comparison cannot in the end override the model's rejection when compared with $\mathcal{T}_2$

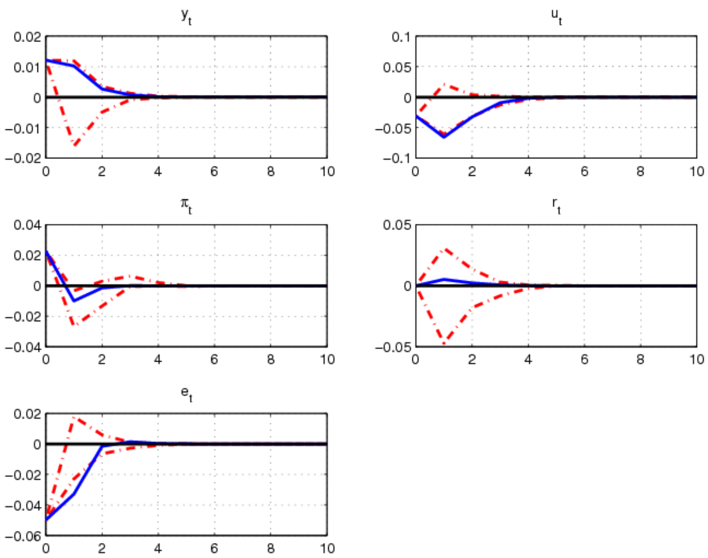

Figure 4: **Exchange Rate Targeting Regime / Nominal Shock**

since the latter representation not only fits the data better but also corresponds strictly to the model's own restrictions which impose separate and distinct regimes. Thus the comparison with $\mathcal{T}_1$ is telling us that the model would fit the dynamic facts better if modified, which is also the message of the comparison with $\mathcal{T}_2$.

4 Conclusion

In this paper we have introduced a new bootstrap method, based on the method of indirect inference, for testing macroeconomic models according to their dynamic performance. The method maintains a separation between the structural (non-linear) model as the null hypothesis and the dynamic time series representation of the data. The model's errors are discovered and used for bootstrapping (after whitening); the resulting pseudo-samples are used to discover the sampling distribution of the dynamic time series model. The test then consists of discovering whether functions of the parameters of the time-series model estimated on the actual data lie within some confidence interval of this distribution. A Wald test statistic or overall M(odel)-metric is developed for this purpose.

We demonstrated the use of the method through an application to the Liverpool Model of the UK over recent postwar data, chosen because it has been successfully estimated by FIML, is approximately micro-founded and has a record of some empirical success dynamically. We tested this model as proposed here and found it to be rejected at the 95% level, though just accepted

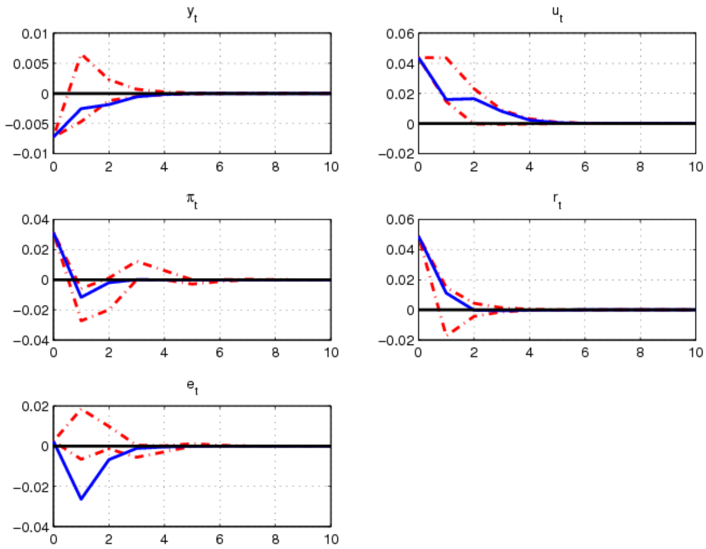

Figure 5: **Inflation Targeting Regime / Supply Shock**

when the data is represented by a VAR with state-varying regimes. This result suggests that with some adjustments, perhaps to the scope of each regime, this model could achieve dynamic acceptance; this use of rejection to produce new and better hypotheses was what Popper (1934) suggested was the way in which science progressed. More generally, we conclude that the method seems to be useful but should be applied to other models in further work to establish its properties more widely.

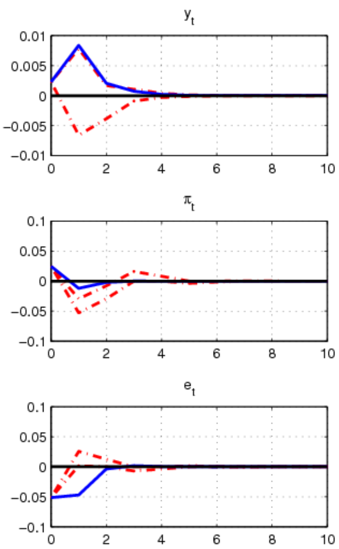

Figure 6: **Inflation Targeting Regime / Nominal Shock**

Table 12: **Distribution of the Autoregressive Parameters of $\mathcal{T}_1$**

Parameters	*Actual Estimates*	*Lower Bound*	*Upper Bound*	*State*
$c_{\Delta y}$	-0.0003	-0.0321	0.0131	IN
$c_{\Delta u}$	-0.0142	-0.0347	0.0351	IN
$c_{\Delta \pi}$	0.0006	-0.0387	0.0173	IN
$c_{\Delta r}$	-0.0015	-0.0252	0.0116	IN
$c_{\Delta e}$	0.0065	-0.0343	0.0178	IN
$A^{\Delta y}_{\Delta y}$	-0.2377	-0.9734	0.9061	IN
$A^{\Delta y}_{\Delta u}$	2.0814	-1.5017	1.5007	OUT
$A^{\Delta y}_{\Delta \pi}$	0.5809	-1.1047	1.2411	IN
$A^{\Delta y}_{\Delta r}$	0.7502	-0.816	0.7364	OUT
$A^{\Delta y}_{\Delta e}$	-0.1129	-1.1481	0.9939	IN
$A^{\Delta u}_{\Delta y}$	0.0762	-0.2589	0.2895	IN
$A^{\Delta u}_{\Delta u}$	0.1478	-0.413	0.5709	IN
$A^{\Delta u}_{\Delta \pi}$	0.0699	-0.2935	0.3651	IN
$A^{\Delta u}_{\Delta r}$	0.1073	-0.1925	0.278	IN
$A^{\Delta u}_{\Delta e}$	0.2129	-0.4088	0.3353	IN
$A^{\Delta \pi}_{\Delta y}$	0.0656	-0.6312	0.589	IN
$A^{\Delta \pi}_{\Delta u}$	0.2451	-1.0365	1.0584	IN
$A^{\Delta \pi}_{\Delta \pi}$	0.0551	-0.6704	0.762	IN
$A^{\Delta \pi}_{\Delta r}$	0.099	-0.6169	0.3988	IN
$A^{\Delta \pi}_{\Delta e}$	0.2492	-0.863	0.7246	IN
$A^{\Delta r}_{\Delta y}$	0.6101	-0.2411	0.2166	OUT
$A^{\Delta r}_{\Delta u}$	-0.5829	-0.3497	0.4128	OUT
$A^{\Delta r}_{\Delta \pi}$	-0.0322	-0.3376	0.2661	IN
$A^{\Delta r}_{\Delta r}$	-0.2724	-0.2999	0.3293	IN
$A^{\Delta r}_{\Delta e}$	0.4387	-0.3468	0.3038	OUT
$A^{\Delta e}_{\Delta y}$	-0.061	-0.6441	0.4252	IN
$A^{\Delta e}_{\Delta u}$	-0.0043	-0.74	0.8845	IN
$A^{\Delta e}_{\Delta \pi}$	-0.1366	-0.5907	0.7798	IN
$A^{\Delta e}_{\Delta r}$	-0.1213	-0.4298	0.5003	IN
$A^{\Delta e}_{\Delta e}$	-0.1555	-0.6315	0.522	IN

Table 13: **Distribution of the Covariance Parameters of $\mathcal{T}_1$**

Covariances	*Actual estimates*	*Lower Bound*	*Upper Bound*	*State*
$\sigma^1_{\Delta y\Delta y}$	0.0015	0.0004	0.1648	IN
$\sigma^1_{\Delta u\Delta y}$	0.0034	-0.0528	0.0755	IN
$\sigma^1_{\Delta \pi\Delta y}$	0.0009	-0.0346	0.0643	IN
$\sigma^1_{\Delta r\Delta y}$	0.0013	-0.0297	0.0251	IN
$\sigma^1_{\Delta e\Delta y}$	0.0007	-0.0133	0.0517	IN
$\sigma^1_{\Delta u\Delta u}$	0.0123	0.003	0.7145	IN
$\sigma^1_{\Delta \pi\Delta u}$	0.0022	-0.0514	0.0594	IN
$\sigma^1_{\Delta r\Delta u}$	0.0028	-0.065	0.0864	IN
$\sigma^1_{\Delta e\Delta u}$	0.0029	-0.0892	0.0318	IN
$\sigma^1_{\Delta \pi\Delta \pi}$	0.001	0.0007	0.2563	IN
$\sigma^1_{\Delta r\Delta \pi}$	0.0008	-0.057	0.0262	IN
$\sigma^1_{\Delta e\Delta \pi}$	0.0004	-0.1038	0.0292	IN
$\sigma^1_{\Delta r\Delta r}$	0.0013	0.0006	0.293	IN
$\sigma^1_{\Delta e\Delta r}$	0.0004	-0.0265	0.0399	IN
$\sigma^1_{\Delta e\Delta e}$	0.0013	0.0005	0.2484	IN
$\sigma^2_{\Delta y\Delta y}$	0.0008	0.0003	0.1238	IN
$\sigma^2_{\Delta u\Delta y}$	0.001	-0.0299	0.0537	IN
$\sigma^2_{\Delta \pi\Delta y}$	0.0005	-0.0125	0.0297	IN
$\sigma^2_{\Delta r\Delta y}$	0.0006	-0.0301	0.0271	IN
$\sigma^2_{\Delta e\Delta y}$	0.0008	-0.0169	0.0382	IN
$\sigma^2_{\Delta u\Delta u}$	0.0022	0.0017	0.1966	IN
$\sigma^2_{\Delta \pi\Delta u}$	0.0003	-0.0405	0.0292	IN
$\sigma^2_{\Delta r\Delta u}$	0.0000	-0.0256	0.0955	IN
$\sigma^2_{\Delta e\Delta u}$	-0.0001	-0.0321	0.0224	IN
$\sigma^2_{\Delta \pi\Delta \pi}$	0.0004	0.0006	0.2037	OUT
$\sigma^2_{\Delta r\Delta \pi}$	0.0004	-0.0462	0.0196	IN
$\sigma^2_{\Delta e\Delta \pi}$	0.0003	-0.0367	0.0135	IN
$\sigma^2_{\Delta r\Delta r}$	0.0005	0.0005	0.187	IN
$\sigma^2_{\Delta e\Delta r}$	0.0007	-0.0308	0.0363	IN
$\sigma^2_{\Delta e\Delta e}$	0.0021	0.0006	0.1805	IN
$\sigma^3_{\Delta y\Delta y}$	0.0009	0.0003	0.1048	IN
$\sigma^3_{\Delta u\Delta y}$	0.0021	-0.0373	0.0258	IN
$\sigma^3_{\Delta \pi\Delta y}$	0.0007	-0.011	0.03	IN
$\sigma^3_{\Delta r\Delta y}$	0.0007	-0.0109	0.0157	IN
$\sigma^3_{\Delta e\Delta y}$	0.0005	-0.012	0.0196	IN
$\sigma^3_{\Delta u\Delta u}$	0.0063	0.002	0.1809	IN
$\sigma^3_{\Delta \pi\Delta u}$	0.002	-0.0406	0.0169	IN
$\sigma^3_{\Delta r\Delta u}$	0.0016	-0.0092	0.0477	IN
$\sigma^3_{\Delta e\Delta u}$	0.0015	-0.0221	0.0201	IN
$\sigma^3_{\Delta \pi\Delta \pi}$	0.0007	0.0006	0.1356	IN
$\sigma^3_{\Delta r\Delta \pi}$	0.0006	-0.0233	0.0058	IN
$\sigma^3_{\Delta e\Delta \pi}$	0.0005	-0.0382	0.0113	IN
$\sigma^3_{\Delta r\Delta r}$	0.0006	0.0005	0.0826	IN
$\sigma^3_{\Delta e\Delta r}$	0.0004	-0.0079	0.0195	IN
$\sigma^3_{\Delta e\Delta e}$	0.0005	0.0004	0.0852	IN

Table 14: **M-metric (Wald statistic) for Model using $\mathcal{T}_1$**

CI	***Bootstrapped Statistic***	***Actual Statistic***
94%	0.032	
95%	**0.0354**	***0.0344***
96%	0.0366	
97%	0.0389	
98%	0.0414	
99%	0.0507	
100%	0.3921	

References

Andrews, M., Minford, P. and Riley, J. (1996), 'On comparing macroeconomic models using forecast encompassing tests', *Oxford Bulletin of Economics and Statistics* **58**, 279–305.

Burnside, C., Eichenbaum, M. and Rebelo, S. (1993), 'Labor hoarding and the business cycle', *Journal of Political Economy* **101**, 245–273.

Canova, F. (2005), *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton.

Christiano, L. and Eichenbaum, M. (1992), 'Current real business cycle theories and aggregate labor market fluctuation', *American Economic Review* **82**, 430–450.

Christiano, L., Eichenbaum, M. and Evans, C. (2005), 'Nominal rigidities and the dynamic effects of a shock to monetary policy', *Journal of Political Economy* **113**, 1–45.

Cogley, T. and Nason, J. (1994), 'Testing the implications of long-run neutrality for monetary business cycle models', *Journal of Applied Econometrics* **9**, S37–S70.

Corradi, V. and Swanson, N. (2005), 'Evaluation of dynamic stochastic general equilibrium models based on distributional comparison of simulated and historical data', *Journal of Econometrics* **forthcoming**.

DeJong, D., Ingram, B. and Whiteman, C. (2000), 'A bayesian approach to dynamic macroeconomics', *Journal of Econometrics* **98**, 203–223.

Del Negro, M. and Schorfheide, F. (2004), 'Priors from general equilibrium models for vars', *International Economic Review* **45**, 643–673.

Diebold, F., Ohanian, L. and Berkowitz, J. (1998), 'Dynamics general equilibrium economies: A framework for comparing models and data', *Review of Economic Studies* **65**, 433–451.

Fair, R. and Taylor, J. (1983), 'Solution and maximum likelihood estimation of dynamic nonlinear rational expectation models', *Econometrica* **51**, 1169–1185.

Fernandez-Villaverde, J. and Rubio-Ramirez, J. (2004), 'Comparing dynamic equilibrium models to data: A bayesian approach', *Journal of Econometrics* **123**, 153–185.

Fève, P. and Langot, F. (1994), 'The rbc models through statistical inference: An application with french data', *Journal of Applied Econometrics* **9**, S11–S37.

Gourieroux, C. and Monfort, A. (1995), *Simulation Based Econometric Methods*, CORE Lectures Series, Louvain-la-Neuve.

Gourieroux, C., Monfort, A. and Renault, E. (1993), 'Indirect inference', *Journal of Applied Econometrics* **8**, 85–118.

Gregory, A. and Smith, G. (1991), 'Calibration as testing: Inference in simulated macro models', *Journal of Business and Economic Statistics* **9**, 293–303.

Gregory, A. and Smith, G. (1993), Calibration in macroeconomics, *in* G. Maddala, ed., 'Handbook of Statistics'. vol. 11, Elsevier, St.Louis, Mo., 703-719.

Krolzig, H. (1997), 'Markov-switching vector autoregressions. modelling, statistical inference and application to business cycle analysis', *Lecture Notes in Economics and Mathematical Systems* **Volume 454**.

Lutkepohl, H. (1993), *Introduction to Multiple Time Series Analysis*, Springer, Berlin.

McCallum, B. and Nelson, E. (1999), 'Nominal income targeting in an open economy optimising mode', *Journal of Monetary Economics* **43**, 553–578.

McCallum, B. and Nelson, E. (2000), 'Monetary policy for an open economy: An alternative framework with optimizing agents and sticky prices', *Oxford Review of Economic Policy* **16**, 74–91.

Meenagh, D., Minford, P., Nowell, E. and Sofat, P. (2008), 'Real exchange rate overshooting rbc style', *Working Paper* **Cardiff**.

Minford, P. (1980), A rational expectations model of the united kingdom under fixed and floating exchange rates, *in* K. Brunner and A. Meltzer, eds, 'On the State of Macroeconomics, Carnegie Rochester Conference Series on Public Policy, 12'. Supplement to the Journal of Monetary Economics.

Minford, P., Meenagh, D. and Webb, B. (2004), 'Britain and emu: Assessing the costs in macroeconomic variability', *The World Economy* **27**, 301–358.

Minford, P. and Webb, B. (2005), 'Estimating large rational expectations models by fiml: A new algorithm with bootstrap confidence limits', *Economic Modelling* **21**, 187–205.

Schorfheide, F. (2000), 'Loss function based evaluation of dsge models', *Journal of Applied Econometrics* **15**, 645–670.

Smets, F. and Wouters, R. (2003), 'An estimated stochastic dsge model of the euro area', *Journal of the European Economic Association* **1**, 1123–1175.

Smith, A. (1993), 'Estimating nonlinear time-series models using simulated vector autoregressions', *Journal of Applied Econometrics* **8**, S63–S84.

Soderlind, P. (1994), 'Cyclical properties of a real business cycle model', *Journal of Applied Econometrics* **9**, S113–S122.

Stinchcombe, M. and White, H. (1992), 'Some measurability results for extremea of random functions over random sets', *Review of Economic Studies* **59**, 495–514.

Theodoridis, K. (2006), Evaluating Dynamic General Equilibrium Models, PhD thesis, Cardiff Business School.

Watson, M. (1993), 'Measures of fit for calibrated models', *Journal of Political Economy* **101**, 1011–1041.

White, H. (1994), *Estimation, Inference and Specification Analysis*, Cambridge University Press, Cambridge.