

How much nominal rigidity is there in the US economy? Testing a New Keynesian DSGE Model using indirect inference*

Vo Phuong Mai Le (Cardiff University)[†]

Patrick Minford (Cardiff University and CEPR)[‡]

Michael Wickens (Cardiff University, University of York and CEPR)[§]

January 18, 2010

Abstract

We evaluate the Smets-Wouters model of the US dynamically using indirect inference with a VAR representation of the main US data series. We find that the New Keynesian SW model is badly rejected by the data's dynamic properties and in particular cannot match the variability of the data. An alternative (New Classical) version of the model with flexible wages and prices and a one-period information lag fares no better. A 'weighted' model (mostly NC but part NK) is able to match the data variability, though it too is rejected overall. Allowing for structural breaks in monetary regime we find a model from 1984 onwards fits fairly well dynamically; this has a high NK weight, suggesting much greater nominal stickiness during the 'great moderation'.

JEL Classification: C12, C32, C52, E1,

Keywords: Bootstrap, US model, DSGE, VAR, New Keynesian, New Classical, indirect inference, Wald statistic, regime change, structural break, great moderation

*We thank Huw Dixon, Laurence Copeland and David Meenagh for their helpful comments. This work was supported by the UK's Economic and Social Research Council under grants RES-165-25-0020 and PTA-026-27-1623. A full set of results can be found in the Annexes to this paper at www.cf.ac.uk/carbs/research/sections/econ/workingpapers.html/papers/E2008_32.pdf

[†]Levp@cf.ac.uk; Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK

[‡]Patrick.minford@btinternet.com; Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK

[§]Cardiff Business School, Cardiff University, Aberconway Building, Colum Drive, Cardiff, CF10 3EU, UK

1 Introduction

In this paper we test a dynamic stochastic general equilibrium (DSGE) model of the US economy on the full sample of post-war data using the method of indirect inference. The aim of the tests is to determine the degree of nominal rigidity in the US economy. The extent of nominal rigidity is a major area of disagreement between economists and we believe — and this is confirmed by our results — that it holds a critical key to the model's dynamic performance. Our modelling framework is based on that of Smets and Wouters (2007). Smets and Wouters (SW) adopt a New Keynesian (NK) model with sticky prices and wages. We compare this with a New Classical (NC) version of their model which has flexible prices and wages and a one-quarter delay for households in receiving macro information. We also consider the possibility that the economy consists of a mixture of the two in which some parts of the economy display nominal rigidities and other parts do not.

To anticipate our results, we find that for the full sample period a hybrid model in which most of the economy enjoys price and wage flexibility but a non-negligible part is subject to nominal contracts, gets closest to matching the data, whereas the NK and NC models are seriously at odds with the data. If, however, we restrict the data to a sub-period from the mid-1980s to the mid-2000s then a model with a high degree of nominal rigidity is then able to match key aspects of the data. Our results suggest that the state-dependency of pricing could dominate its time-dependency for the bulk of the post-war period but that during the later period of the 'great moderation', when the economy was more stable, time-dependency could have dominated.

The SW model contains a full range of structural shocks and nominal and real frictions, and the model is estimated with Bayesian methods. They consider their model satisfactory in a variety of ways. For example, it can compete with standard VAR and BVAR models in forecasting the main US macro variables at business cycle frequencies. In this paper we focus on the model's dynamic performance within the sample, using a new evaluation procedure based on indirect inference. This exploits the properties of the model's error processes through bootstrap simulations and analyses whether the simulated data from the structural model can explain the actual data when both are represented by the dynamic behaviour of an auxiliary model. Our test, which is a form of Wald statistic, focuses on the overall capacity of the model to fit the data's dynamic performance.

The idea is to use indirect inference to test the structural model. Indirect inference has been widely used in the estimation of structural models, see Smith (1993), Gregory and Smith (1991,1993), Gourieroux et al. (1993), Gourieroux and Monfort (1995) and Canova (2005). Here we make a different use of indirect inference as our aim is to evaluate an already estimated or calibrated structural model. The common element is the use of an auxiliary time series model. In estimation the parameters of the structural model are chosen so that when this model is simulated it generates estimates of the auxiliary model similar to those obtained from actual data. The optimal choice of parameters for the structural model are those that minimise the distance between a given function of the two sets of estimated coefficients of the auxiliary model. Common choices of this function are the actual coefficients, the scores or the impulse response functions. In model evaluation the parameters of the structural model are taken as given. The aim is to compare the performance of the auxiliary model estimated on simulated data derived from the given estimates of a structural model - which is taken as a true model of the economy, the null hypothesis - with the performance of the auxiliary model when estimated from actual data. If the structural model is correct then its predictions about the impulse responses, moments and time series properties of the data should statistically match those based on actual data. The comparison is based on the distributions of the two sets of parameter estimates of the auxiliary model, or of functions of these estimates.

In other words, the testing procedure involves first constructing the errors derived from the previously estimated structural model and the actual data. These errors are then bootstrapped and used to generate for each bootstrap new data based on the structural model. An auxiliary time-series model is then fitted to each set of data and the sampling distribution of the coefficients of the auxiliary time series model is obtained from these estimates of the auxiliary model. A Wald statistic is computed to determine whether functions of the parameters of the time series model estimated on the actual data lie in some confidence interval implied by this sampling distribution.

This paper builds on Le (2008), who examined the ability of the calibrated model of Canzoneri, Cumby and Diba (2004) to fit the US data. This is a very simple New Keynesian model, designed to investigate

certain policy issues by simulation; it contains Calvo contract assumptions in labour and product markets and investment adjustment costs but no other rigidities. It is rejected by the Wald-statistic and other standard ways of measuring the goodness of fit. Furthermore, whatever assumptions were added to the model about indexation and whatever assumptions were made about the degree of nominal rigidity, including moving to a New Classical version, made no impact on the data-acceptability of the model; all versions were rejected at high levels of significance. In particular, the model did not have sufficient sources of lagged reaction to fit the data well.

The SW model for the US economy has many of the features of the Canzoneri et al. model, including Calvo contracts in product and labour markets. But it also embeds backward-looking indexation and real rigidities other than and in addition to investment adjustment costs — both habit formation in consumption and variable capital-utilisation with its own adjustment costs. These are features that New Keynesian economists advocate in order to replicate the persistence of inflation and output and supposedly hump shaped responses to monetary shocks — e.g. Christiano et al (2007).

We find in what follows that neither NK nor NC models can be used to represent the U.S. economy on their own, because both of them fail to satisfy the most basic measures of fitness. However, when a limited degree of nominal rigidity is embedded in the NC model the resulting ‘mixed’ model fits the data better. This means that, although nominal rigidity is relevant in modelling the economy, its role is not as important as New Keynesian economists propose. Real rigidities on the other hand are highly necessary to replicate the data.

We do not directly consider micro data on price-setting such as that examined by Bils and Klenow (2004), but we recognise that work in this field is continuing in order to establish how much nominal rigidity is present. We also note that recently Gertler and Leahy (2008) have suggested that state-contingent (Ss) pricing could account both the micro evidence on sticky prices as well as macro price behaviour and that this could look quite like price-flexibility.

The paper is organised as follows. In section 2 we describe the Smets-Wouters model and summarise their main findings. We explain the testing procedure in detail in section 3. In section 4 first we compare the NK and NC models, and then we compare these with the hybrid model. In section 5 we examine whether changes in monetary regimes are a possible source of misspecification. We summarise our conclusions in section 6.

2 The Smets-Wouters model of the US

Smets and Wouters (2007) developed a New Keynesian model and estimated this model on US data covering the period 1966Q1–2004Q4 using Bayesian methods. The model features many nominal and real frictions that create hump-shaped responses of aggregate demand to shocks. The model’s dynamics are driven by seven orthogonal structural shocks: total factor productivity shocks, risk premium shocks, investment-specific technology shocks, wage mark-up shocks, price mark-up shocks, exogenous spending shocks and monetary policy shocks.

Their model is based on Smets and Wouters (2003) which was estimated on EU data. In the US version there are a few differences. First, the number of structural shocks is reduced from ten to seven. Second, the Dixit-Stiglitz aggregator in the goods and labour markets is replaced by the aggregator developed by Kimball (1995) where the demand elasticity of differentiated goods and labour depends on their relative price. Third, the model features a deterministic growth rate driven by labour-augmenting technological progress which is assumed in order to use the original data without having to detrend them. Their model is estimated by Bayesian methods which combine calibrated parameters with sample information.

SW report that the estimated model fits the US data quite well. This is verified by comparing the marginal likelihood of out-of-sample predictions of the model with Bayesian VAR models. Price and wage rigidities are important in explaining the data but the indexation is not. They find that demand shocks, such as those to the risk premium and to exogenous spending, and investment specific technology shocks explain a significant fraction of the short-run forecast variance in output, but wage mark-up and productivity shocks contribute little to explaining output variation in the medium to long run. They also confirm that productivity shocks have a significant short-run negative impact on hours worked. Inflation developments are mostly driven by the price mark-up shocks in the short run and wage mark-up shocks in the long run.

The model can capture the cross correlation between output and inflation at business cycle frequencies. As an ultimate check of the model's performance, they estimate the model for two subsamples: the "Great Inflation" period from 1966Q2 to 1979Q2 and the "Great Moderation" period from 1984Q1 to 2004Q4, and find that most of the structural parameters are stable over those periods except for the fall in the standard deviation of the productivity, monetary policy and price mark-up shocks, which reflect the decrease in output growth and inflation volatility, and for the fall in the monetary policy response to output developments in the second subsample.

We note that all of these exercises are carried out in a Bayesian framework and that at no stage is the specification of the model tested. In effect, the Bayesian procedure just improves the fit compared with using only calibrated parameters by employing additional information from the data. Nonetheless, the original calibration dominates the final coefficient estimates if the priors are too tight, see for example Del Negro and Schorfheide (2007). A full test of the model in our sense does not arise under a Bayesian procedure because the information in the priors is regarded as fixed and known. One reason why we are carrying out such tests is that we do not accept these priors uncritically; in particular, we are concerned about the extent of nominal rigidity assumed in the priors.

3 Model evaluation by indirect inference

Our aim is to evaluate an already estimated or calibrated (DSGE) macroeconomic model by indirect inference. By evaluate we mean carry out classical statistical inference on a previously estimated or calibrated model. This is related to, but is different from, estimating a macroeconomic model by indirect inference. The common feature is the use of an auxiliary model in addition to the structural macroeconomic model. Before considering model evaluation by indirect inference, to set the scene and establish notation, first we discuss estimation by indirect inference.

3.1 Estimation

Estimation by indirect inference chooses the parameters of the macroeconomic model so that when this model is simulated it generates estimates of the auxiliary model similar to those obtained from the observed data. The optimal choice of parameters for the macroeconomic model are those that minimize the distance between a given function of the two sets of estimated coefficients of the auxiliary model. Common choices of this function are (i) the actual coefficients, (ii) the scores, and (iii) the impulse response functions. In effect, estimation by indirect inference gives the optimal calibration.

Suppose that y_t is an $m \times 1$ vector of observed data, $t = 1, \dots, T$, $x_t(\theta)$ is an $m \times 1$ vector of simulated time series generated from the structural macroeconomic model, θ is a $k \times 1$ vector of the parameters of the macroeconomic model and $x_t(\theta)$ and y_t are assumed to be stationary and ergodic. The auxiliary model is $f[y_t, \alpha]$. 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, i.e.

$$f[x_t(\theta_0), \alpha] = f[y_t, \alpha]$$

where α is the vector of parameters of the auxiliary model.

The likelihood function for the auxiliary model defined for the observed data $\{y_t\}_{t=1}^T$ is

$$\mathcal{L}_T(y_t; \alpha) = \sum_{t=1}^T \log f[y_t, \alpha].$$

The maximum likelihood estimator of α is then

$$a_T = \arg \max_{\alpha} \mathcal{L}_T(y_t; \alpha).$$

The corresponding likelihood function based on the simulated data $\{x_t(\theta)\}_{s=1}^S$ is

$$\mathcal{L}_S[x_t(\theta); \alpha] = \sum_{t=1}^S \log f[x_t(\theta), \alpha]$$

with

$$a_S(\theta) = \arg \max_{\alpha} \mathcal{L}_S[x_t(\theta); \alpha].$$

The simulated quasi-maximum likelihood estimator (SQMLE) of θ is

$$\theta_{T,S} = \arg \max_{\theta} \mathcal{L}_T[y_t; \alpha_S(\theta)].$$

This is the value of θ that produces a value of α that maximises the likelihood function using the observed data. We suppose that the observed and the simulated data are such that this value of α satisfies

$$\text{plim } a_T = \text{plim } a_S(\theta) = \alpha,$$

hence the assumption that $x_t(\theta)$ and y_t are stationary and ergodic, see Canova (2005). It can then be shown that

$$\begin{aligned} T^{1/2}(a_S(\theta) - \alpha) &\rightarrow N[0, \Omega(\theta)] \\ \Omega(\theta) &= E\left[-\frac{\partial^2 \mathcal{L}[\alpha(\theta)]}{\partial \alpha^2}\right]^{-1} E\left[\frac{\partial \mathcal{L}[\alpha(\theta)]}{\partial \alpha} \frac{\partial \mathcal{L}[\alpha(\theta)]'}{\partial \alpha}\right] E\left[-\frac{\partial^2 \mathcal{L}[\alpha(\theta)]}{\partial \alpha^2}\right]^{-1}. \end{aligned}$$

The covariance matrix can be obtained either analytically or by bootstrapping the simulations.

An alternative to the SQMLE is the extended method of simulated moments estimator (EMSME). This is obtained as follows. Consider the continuous $p \times 1$ vector of functions $g(a_T)$ and $g(\alpha_S(\theta))$ which could, for example, be moments or scores, and let $G_T(a_T) = \frac{1}{T} \sum_{t=1}^T g(a_T)$ and $G_S(\alpha_S(\theta)) = \frac{1}{S} \sum_{s=1}^S g(\alpha_S(\theta))$. We require that $a_T \rightarrow \alpha_S$ in probability and that $G_T(a_T) \rightarrow G_S(\alpha_S(\theta))$ in probability for each θ . The EMSME is

$$\theta_{T,S} = \arg \min_{\theta} [G_T(a_T) - G_S(\alpha_S(\theta))] W(\theta) [G_T(a_T) - G_S(\alpha_S(\theta))].$$

3.2 Model evaluation

The parameters of the macroeconomic model and their distributions are taken as given — either estimated or calibrated. The aim is to compare the performance of the auxiliary model based on observed data with its performance based on simulations of the macroeconomic model derived by bootstrapping its structural disturbances. From these simulations we may obtain the joint distribution of the parameters of the auxiliary model and use this to perform a Wald test. This tests whether the estimates of the auxiliary model based on actual data could have come from the particular realisation of the structural model. We choose a VAR as the auxiliary model and base our test on a function of the VAR coefficients (augmented by the data variances, as a check for matching variability). We use a VAR(1) on a limited number of key variables. By raising the order of the VAR and increasing the number of variables, the stringency of the overall test of the model is increased. As we find that the structural model is already rejected by a VAR(1), we do not proceed to a more stringent test based on a higher order VAR.

Non-rejection of the null hypothesis is taken to indicate that dynamic behaviour of the macroeconomic model is not significantly different from that of the observed data. Rejection is taken to imply that the macroeconomic model is incorrectly specified. Comparison of the impulse response functions of the observed and simulated data should reveal in what respects the macroeconomic model fails to capture the auxiliary model.

The Wald test statistic is obtained as follows. 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 $c \geq 1$. If $\hat{\theta}$ is the estimated or calibrated value of θ then the null hypothesis can be expressed as $H_0 : \hat{\theta} \rightarrow \theta_0$. Consider again the continuous $p \times 1$ vector of functions $g(a_T)$, $g(\alpha_S(\theta))$, $G_T(a_T) = \frac{1}{T} \sum_{t=1}^T g(a_T)$ and $G_S(\alpha_S(\theta)) = \frac{1}{S} \sum_{s=1}^S g(\alpha_S(\theta))$. The functions $g(\cdot)$ may be impulse response functions. Given an auxiliary model and a function of its parameters, our test statistic for evaluating the macroeconomic model is based on the distribution of $G_T(a_T) - G_S(\alpha_S(\hat{\theta}))$. The resulting Wald statistic may be written as

$$[G_T(a_T) - G_S(\alpha_S(\hat{\theta}))]' W(\hat{\theta}) [G_T(a_T) - G_S(\alpha_S(\hat{\theta}))]$$

where the estimate of the optimal weighting matrix is

$$W(\hat{\theta}) = \left\{ \left[\frac{\partial G(\alpha(\hat{\theta}))}{\partial \alpha} \right] \Omega(\hat{\theta}) \left[\frac{\partial G(\alpha(\hat{\theta}))}{\partial \alpha} \right]' \right\}^{-1}$$

We obtain the distribution of $G_T(a_T) - G_S(\alpha_S(\hat{\theta}))$ and the Wald statistic using the bootstrap.

The following steps summarise our implementation of the Wald test by bootstrapping:

Step 1: Estimate the errors of the economic model conditional on the observed data and $\hat{\theta}$.

Estimate the DSGE macroeconomic model's structural the errors ε_t given $\hat{\theta}$ and the observed data. The number of independent structural errors is taken to be less than or equal to the number of endogenous variables. The errors are not assumed to be normally distributed.

Step 2: Estimate the empirical distribution of the structural errors

On the null hypothesis the $\{\varepsilon_t\}_{t=1}^T$ errors are omitted variables. Their empirical distribution is assumed to be given by these structural errors. The simulated disturbances are drawn from these errors. In some DSGE models the structural errors are assumed to be generated by autoregressive processes which under our method we need to estimate. This is the case with the SW model. The model is bootstrapped (drawing these disturbances by time vector to preserve any simultaneity between them) and solved using Dynare (Juillard, 2001).

Step 3: Compute the Wald statistic

We choose the function of the auxiliary model's parameters to be the VAR coefficients themselves rather than a multi-valued function of them such as the impulse response functions (IRFs). Hence

$$g(a_T) - g(\alpha_S(\theta)) = a_T - \alpha_S(\theta)$$

and so

$$G_T(a_T) - G_S(\alpha_S(\hat{\theta})) = a_T - \alpha_S(\hat{\theta})$$

The distribution of $a_T - \alpha_S(\hat{\theta})$ and its covariance matrix $W(\hat{\theta})^{-1}$ are estimated by bootstrapping $\alpha_S(\hat{\theta})$. Thus we use the appropriate small-sample distribution rather than the asymptotic distribution that emerges from analytic methods.

The bootstrapping proceeds by drawing N bootstrap samples of the structural model, and estimating the auxiliary VAR on each, thus obtaining N $\alpha_S(\hat{\theta})$. This set of vectors represents the sampling variation implied by the structural model, enabling its mean, covariance matrix and confidence bounds to be calculated directly. N is generally set to 1000.

We can now compute the properties of the model and compare them with those of the data; in particular, we examine the model's ability to encompass the variances of the data. Assuming the model can do so, we go on to compute the bootstrap Wald statistic $[a_T - \alpha_S(\hat{\theta})]'W(\hat{\theta})^{-1}[a_T - \alpha_S(\hat{\theta})]$. Figure 1 shows, for just two parameters in the auxiliary equation, the distribution of the statistic and an example of the statistic for two cases — one with a diagonal covariance matrix and one with non-zero covariances. One can think of estimation via indirect inference as changing the parameters of the structural model, thus changing the implied distribution, so as to push the observed data point as far into the centre of the distribution as possible. The test however takes the structural parameters as given and merely notes the position of the observed data point in the distribution.

In addition to our basic Wald statistic we consider a number of related Wald statistics. The basic Wald test is based on the full joint distribution of the VAR coefficients as implied by their full covariance matrix, as in the second panel of Figure 1. We refer to this as the Full Wald test; it checks whether the VAR data-based coefficients lie within the DSGE model's implied joint distribution and is a test of the DSGE model's specification in absolute terms. We use the Mahalanobis Distance based on the same joint distribution, normalised as a t-statistic, as an overall measure of closeness between the model and the data- in effect this conveys the same information as in the Wald test but in the form of a t-value.

A second Wald test, which we refer to as a 'Directed Wald statistic', focuses on specific features of the structural model. Here we seek to know how well a particular variable or limited set of variables is modelled and we use the corresponding auxiliary equations for these variables in the VAR as the basis of our test. For example, we may wish to know how well the model can reproduce the behaviour of US and EU output by creating a Wald statistic based on the VAR equation for the two outputs alone.

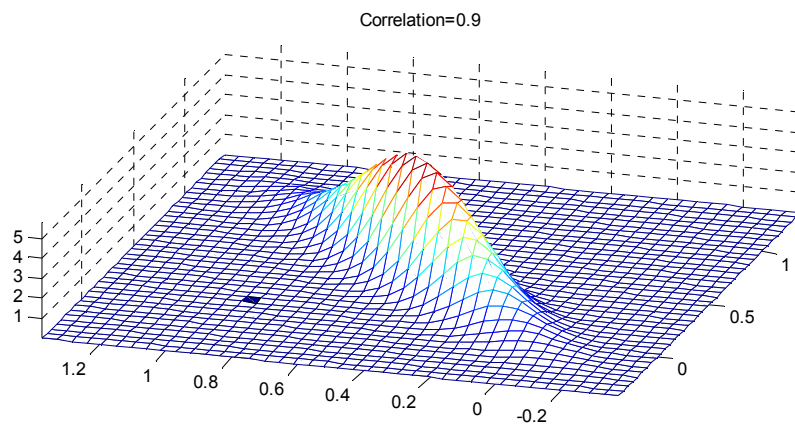
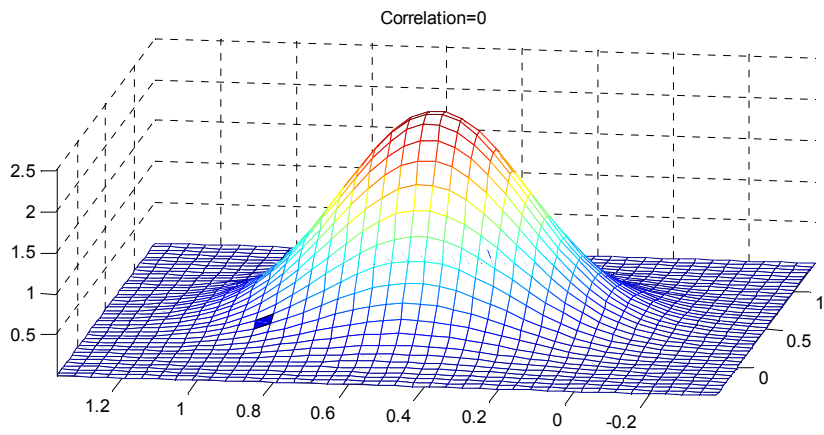


Figure 1: Bivariate Normal Distributions (0.1, 0.9 shaded) with correlation of 0 and 0.9.

Filter/Walds	Variances	Coefficients	Variances and Coefficients
Original	97	100	100
Differenced	100	100	100
HP	100	100	100
Linear Detrend	100	100	100

Table 1: The Wald statistics for each filter on SW’s original model

We may also use this Directed Wald test to determine how well the structural model captures the effects of a particular set of shocks. To do this we create the joint distribution of the IRFs for these shocks alone. For example, to determine how well the model deals with supply shocks, we construct the joint distribution of the IRFs for the supply shocks and calculate a Wald statistic for this. Even if a model is mis-specified overall, through these Directed Wald tests we can say whether it is well-specified enough to deal with specific aspects of economic behaviour.

Finally we should note that we also look at the usual battery of diagnostic statistics for these models, such as ability to match (i.e. embrace within 95% limits) data variances, cross-correlations, and VAR-based IRFs. We attach particular importance to the ability to match data variances, arguing that a failure in this dimension is essentially terminal; for this reason we include the data variances in the full Wald statistic. It should be noted that the cross-correlations (other than contemporaneous) and the IRFs are all derived from the VAR coefficients; hence our focus on these rather than the many relationships that can be derived from them. This makes our procedure in many ways quite traditional; our Wald statistics are our main innovation, but these largely summarise the results of these more traditional measures.

We are implicitly assuming that the auxiliary model can distinguish between different structural models. This has been challenged recently by Canova and Sala (2009). They argue that the identification of different DSGE models is ‘weak’ and may give rise to the same VAR. They explain the point analytically by taking a 3-equation (IS-Phillips Curve-Taylor Rule) reduction of a DSGE model, deriving its VAR representation, and pointing out that several parameters cannot be identified. As they note, however, their example is rigged in particular ways. One is that the shocks in the three equations are i.i.d. and there are no lagged endogenous variables (either from adjustment costs or indexation). In DSGE models like the SW, however, shocks are generally autocorrelated and lagged endogenous variables enter widely. As a result, DSGE models like that of SW are substantially over-identified through the rational expectations mechanism and changes in its parameters imply quite different simulation properties. This is illustrated in our results which attempt to distinguish between DSGE models according to the ‘distance’ of their implied VAR from a data-generated VAR - a related approach is due to Del Negro et al (2006). We find large changes in the variances implied by the model as its degree of rigidity is changed. These variances can be thought of as providing the elements in the distance function and they show how much it changes as rigidity changes. We also find that the Mahalanobis Distance of the model varies with modest changes in model specification. This measure takes account of the joint distribution of all the criterion parameters and so sensitively reflects the model’s complete specification.

4 Testing the SW Model using the method of indirect inference

We apply the proposed testing procedure to this model for the period of 1947Q1–2004Q4. To do so we need to choose a filtering method to stationarise the data. We looked at several methods and report basic results on all of them in the Annexes. In all the model used by SW was rejected by the Wald outright when data variances and VAR coefficients were included in it.

These filters were: SW’s own filter (log differencing of all variables except inflation, log of hours worked and interest rates which are left in levels- Annex A); differencing all variables (log differencing as SW but differencing the remaining three- Annex B); an HP filter (Annex C); and linear detrending as followed by SW for their EU work (SW, 2003)- Annex D. In terms of the models’ success in fitting the data in the widest sense there is not much to choose between these filters, as we have noted. If we consider the model’s ability to replicate the data variances alone, again all are rejected.

We therefore decided that, in order to carry out our more detailed investigations of the model, we would

pick the filter that extracted the least information from the raw data- linear detrending. It turned out that this filter was adequate to generate stationarity for all the data. The results reported in the body of the paper that follows all use this filter.

The VAR estimation is performed with five main observable variables: output, investment, consumption, the quarterly interest rate and the quarterly inflation rate; and we use a VAR of order one. The more variables are added to the VAR and the higher its order, the more detailed properties that the model must match and so the higher is the threshold of the test. As we will see below a VAR(1) on this small list of variables is testing enough.

4.1 Evaluating the SW model using SW's own assumed error properties

First we test the original SW model using their Bayesian estimate means for the error variances and autoregressive coefficients. The model is rejected. This is not entirely a surprise since the Bayesian method only updates priors with the data and does not test the model against the dynamics of the data.

The model is rejected with the Full Wald test statistic of 100; its normalised Mahalanobis Distance is 3.4, indicating that the data's dynamic properties are fairly far from the model's. This can be explained by a large number (9 out of 25) of the VAR's parameters that lie outside the 95% model bounds. Based on the their t-stats, some of these coefficients lie a long way outside the confidence intervals; note in particular the partial autocorrelations of consumption and inflation where the model's bounds lie higher than the VAR estimate. One could interpret this as excessive inflation and consumption persistence in the model. Furthermore, the IRFs of the VAR (when identified by the model) frequently lie modestly outside the model bounds (from the model bootstrap distribution of VAR coefficients).

Futhermore the data variances (the bottom 5 entries in Table 2) for nominal variables are too low compared with the data: for interest rates the data variance lies outside the model bounds while for inflation it lies just on the top 95% bound. Overall, therefore, even with SW's own assumed error properties their model is badly out of line with the data. However, as we see in the next section, these properties are by no means the same as those implied by the data under the null that the model holds.

VAR coeffs	Actual Estimate	Lower Bound	Upper Bound	T-stats
A_Y^Y	0.99908	0.71104	0.96272	2.02349
A_Y^R	0.01503	-0.00557	0.04018	0.07322
A_Y^π	-0.00417	-0.0068	0.0673	-1.41540
A_Y^C	0.10174	-0.07815	0.02091	5.03459
A_Y^I	0.22591	-0.27355	0.18519	2.34051
A_R^Y	-0.64529	-1.28857	-0.40445	0.84625
A_R^R	0.85001	0.66138	0.86763	1.60632
A_R^π	0.15154	-0.11021	0.18262	1.56321
A_R^C	-0.5553	-0.83083	-0.2264	-0.16999
A_R^I	-1.7064	-2.33113	0.39231	-1.14775
A_π^Y	0.11612	-0.44029	0.32551	0.93503
A_π^R	0.02374	0.07066	0.26195	-2.89958
A_π^π	0.59496	0.59853	0.85809	-2.05169
A_π^C	-0.38833	-0.56528	-0.04657	-0.63310
A_π^I	-0.25917	-1.90858	0.40689	0.83059
A_C^Y	-0.08009	-0.12788	0.08505	-0.74255
A_C^R	-0.02553	-0.03697	0.00303	-0.98493
A_C^π	0.0121	-0.04785	0.01597	1.56170
A_C^C	0.78488	0.85948	0.95736	-5.18686
A_C^I	-0.4296	-0.36543	0.08677	-2.66676
A_I^Y	0.02034	0.01692	0.08499	-1.53642
A_I^R	0.01022	-0.00484	0.00905	2.21138
A_I^π	0.01159	-0.01534	0.00714	2.63929
A_I^C	0.01957	0.01241	0.04757	-1.12305
A_I^I	1.02924	0.94769	1.08301	0.38801
σ_Y^2	18.32858	8.71183	47.42615	-0.32176
σ_R^2	0.65276	0.19035	0.56812	3.22089
σ_π^2	0.44451	0.18584	0.46733	1.96505
σ_C^2	10.3888	6.19987	45.31804	-0.74834
σ_I^2	71.79914	65.12685	269.8422	-1.23001
Wald stat	100	Mah. Normalised Distance 3.4		

Table 2: VAR Parameters, data variances and Model Bootstrap Bounds of the SW Model with SW's error properties

4.2 Evaluating the SW model using actual errors

So far we have supplied the SW model with essentially imaginary error properties, chosen by assumption. We now turn to the actual errors derived from estimation on the observed data. We estimate the model's structural errors, that is, the residual in each structural equation is given by the actual data and the expected variables in it. For this we followed a procedure of robust estimation of the structural residuals along the lines suggested by McCallum (1976) and Wickens (1982) under which the expectations on the right hand side of each equation are generated by an instrumental variable regression that is implied by the model. The instruments chosen are the lagged values of the endogenous variables. Thus, in effect, the generated expectations used in deriving the residuals are the predictions of the data-estimated VAR.

Seven behavioural residuals are estimated by this means: consumption, investment, productivity, monetary policy, wage- and price-setting, and one exogenous process, government spending, which enters the goods market clearing condition. These residuals are shown in Figure 2 a.

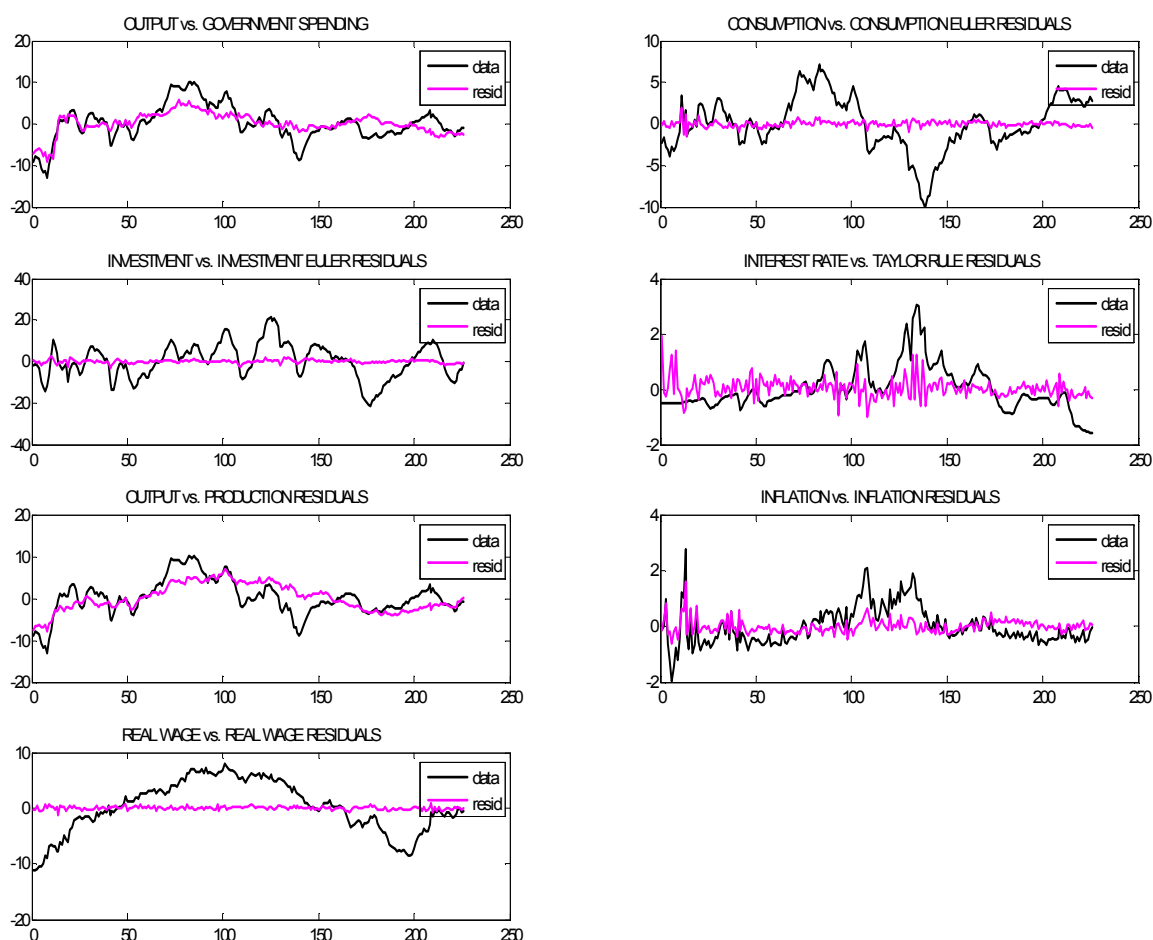


Figure 2: Single Equation Errors from SWNK model

We proceed as though five of these residuals follow an $AR(1)$ and the price and wage residuals follow $ARMA(1, 1)$ processes. The standard deviations of the estimated error innovations are in all cases larger

than those assumed by SW; in the case of investment and the price mark-up they are nearly twice as large (see Table 3). Furthermore, the actual preference, investment and monetary shocks exhibit markedly less persistence than SW assumed. Hence though the properties of the residuals estimated from the data are recognisably similar to those assumed by SW, there are material differences whose effects we go on to investigate in our subsequent bootstrap exercise. We use a vector bootstrap to preserve any dependence between the structural innovations.

	Government Spending	Pref	Inv	Mon	Prod	Price Mark-up	Wage Mark-up
SW stdev	0.53	0.23	0.45	0.24	0.45	0.14	0.24
Data stdev	0.673	0.371	0.704	0.344	0.553	0.239	0.311
SW $AR(1)$	0.97	0.22	0.71	0.15	0.95	0.89	0.96
SW $MA(1)$	0.52					-0.69	-0.84
Estimated $AR(1)$	0.944	-0.064	0.530	-0.062	0.971	0.925	0.915
Estimated $MA(1)$	0.553					-0.709	-0.848

Table 3: Standard deviations of innovations and coefficients of shocks (actual vs. assumed)

The model again fails to capture the scale of the nominal data variances (Table 4, last 5 entries); for the interest rate the data variance is now roughly double the model's upper bound while for inflation it remains around the model upper bound. The results for the VAR coefficients are also reported in Table 4. Based on the Full Wald Statistic, for all of the VAR coefficients and data variances, the model is strongly rejected on at the 5% level. Seven of the VAR coefficients lie outside their 95% bounds, besides the interest data variance. The model's Mahalanobis Distance is 3.6; notice that this is not much different from the SW model using their assumed error properties, so that there is not much change from using the residuals implied by the data.

4.3 Evaluating the New Classical model using actual errors

Next we consider the New Classical version of the SW model proposed above. The results are poor. The main problem is the model's massive overprediction of inflation variance (3rd last entry, Table 5). This occurred regardless of variations in the Taylor Rule; we adopted the NK rule except for setting potential output, y_t^p , to a constant. For example a larger inflation reaction causes the interest rate variance to blow up but without bringing the inflation variance down sufficiently. Thus the model fails on the basic preliminary test of data variance matching.

The model's Full Wald statistic is again 100. Besides the model's overprediction of the inflation variance, out of 25 VAR coefficients, 13 lie outside their 95% bounds. The model wrongly predicts all the partial autocorrelation coefficients, except for that of investment. Of the 13 coefficients that do not fit, five are related to the inflation rate. Further, the cross effects from the main macroeconomic variables to the interest, inflation rates and consumption are badly predicted. The cross-effect from inflation to interest rates in the model is negative; theoretically the interest rate should react to offset a rise in the inflation rate. The Mahalanobis Distance is 7.1 which is considerably worse than for the New Keynesian version of the SW model.

The model's IRFs also perform poorly (see Annex). The dominant shocks on real variables are productivity and labour supply shocks, and on nominal variables are preference, monetary, productivity and labour supply shocks. The responses of all the variables to these shocks lie outside the model 95% bounds. Furthermore, the model fails to replicate the cross-correlations of many of the main macroeconomic variables; it underpredicts the autocorrelations of interest and inflation rates, and their cross-correlations with output; it overpredicts the effect of investment on future output; due to excessive inflation variation, it fails to replicate the correlation between inflation and output.

Overall therefore the New Classical version of the original SW model also fails to match the data in quite serious ways.

4.4 Evaluating a hybrid model: a weighted combination of New Keynesian and New Classical models

We have analysed two rather different macroeconomic models with a view to understand the mechanisms behind each of them. The NK model is highly rigid with Calvo price and wage settings, while the NC is a flexible wage/price model with only a simple one-period information delay for labour suppliers.

In SW's NK model, because capacity utilisation is fairly flexible, output is substantially affected by shocks to demand and this in turn — via the Phillips Curve — moves inflation and then — via the Taylor Rule — interest rates. Supply shocks can affect demand directly (e.g. productivity shocks change the return on capital and so affect investment) and also play a role as 'cost-push' inflation shocks (e.g. price/wage mark-up shocks). Persistent shocks to demand raise 'Q' persistently and produce an 'investment boom' which, via demand effects, reinforces itself. Thus the model acts as a 'multiplier/accelerator' of shocks both on the demand and the supply side.

In the NC model an inelastic labour supply causes output variation to be dominated by supply shocks (productivity and labour supply) while investment and consumption to react to output in a standard RBC manner. These reactions, together with demand shocks, create market-clearing movements in real interest rates and — via the Taylor Rule — in inflation. Supply shocks are prime movers of all variables in the NC model, while demand shocks add to the variability of nominal variables. In order to mimic real variability and persistence, suitably sized and persistent supply shocks are needed. But to mimic the limited variability

VAR coeffs	Actual Estimate	Lower Bound	Upper Bound	T-stats
A_Y^Y	0.99908	0.75267	1.00004	1.52155
A_Y^R	0.01503	0.00444	0.05065	-0.85584
A_Y^π	-0.00417	-0.00459	0.06545	-1.69669
A_Y^C	0.10174	-0.02606	0.07359	2.97157
A_Y^I	0.22591	-0.19814	0.26058	1.59418
A_R^Y	-0.64529	-1.06427	-0.27522	0.12295
A_R^R	0.85001	0.53595	0.73313	3.90309
A_R^π	0.15154	-0.13061	0.15756	1.94297
A_R^C	-0.55553	-0.84942	-0.40004	0.60623
A_R^I	-1.7064	-1.80976	0.3887	-1.85185
A_π^Y	0.11612	-0.55029	0.07888	2.23055
A_π^R	0.02374	0.11413	0.27924	-4.05015
A_π^π	0.59496	0.47347	0.72517	-0.16778
A_π^C	-0.38833	-0.40915	-0.0401	-1.78108
A_π^I	-0.25917	-2.2204	-0.48423	2.53019
A_C^Y	-0.08009	-0.15347	0.11375	-0.78433
A_C^R	-0.02553	-0.06976	-0.01127	0.87955
A_C^π	0.0121	-0.06965	0.02018	1.70294
A_C^C	0.78488	0.8106	0.94474	-2.99327
A_C^I	-0.4296	-0.3507	0.1845	-2.51936
A_I^Y	0.02034	-0.00146	0.06908	-0.48864
A_I^R	0.01022	-0.00373	0.01252	1.44447
A_I^π	0.01159	-0.01151	0.01217	1.76852
A_I^C	0.01957	0.00317	0.03533	-0.02989
A_I^I	1.02924	0.90238	1.03372	1.86272
σ_Y^2	18.32858	9.67374	45.56006	-0.41926
σ_R^2	0.65276	0.16837	0.37665	7.52511
σ_π^2	0.44451	0.22269	0.47431	1.84365
σ_C^2	10.3888	4.62427	35.15967	-0.46506
σ_I^2	71.79914	63.05612	258.1966	-1.24068
Wald stat	100	Mah. Distance (Normalised)		3.6

Table 4: VAR Parameters, data variances and Model Bootstrap Bounds of the SW Model with Estimated Coefficients

VAR coeffs	Actual Estimate	Lower Bound	Upper Bound	T-stats
A_Y^Y	0.99908	0.75964	0.98742	1.81743
A_Y^R	0.01503	-0.0305	0.06499	-0.02425
A_Y^π	-0.00417	-0.03652	0.25225	-1.23410
A_Y^C	0.10174	-0.05527	0.07763	2.72888
A_Y^I	0.22591	-0.18755	0.24126	1.84048
A_R^Y	-0.64529	-0.77893	-0.00061	-1.29619
A_R^R	0.85001	0.16679	0.55328	4.77475
A_R^π	0.15154	-1.70697	-0.51682	4.01721
A_R^C	-0.5553	-0.49638	0.04422	-2.35582
A_R^I	-1.7064	-1.53308	0.39447	-2.34629
A_π^Y	0.11612	-0.14849	0.09029	2.45278
A_π^R	0.02374	0.01012	0.12289	-1.43555
A_π^π	0.59496	0.10313	0.43618	3.75306
A_π^C	-0.38833	-0.15269	0.00878	-7.46875
A_π^I	-0.25917	-0.43894	0.13493	-0.79497
A_C^Y	-0.08009	-0.11079	0.13547	-1.25499
A_C^R	-0.02553	-0.14256	-0.02468	1.75061
A_C^π	0.0121	-0.42415	-0.06702	2.51545
A_C^C	0.78488	0.84915	1.00249	-3.95797
A_C^I	-0.4296	-0.30533	0.16651	-2.94383
A_I^Y	0.02034	-0.00818	0.06222	-0.31562
A_I^R	0.01022	-0.00083	0.03162	-0.58207
A_I^π	0.01159	-0.02859	0.07379	-0.40644
A_I^C	0.01957	-0.00999	0.03741	0.44958
A_I^I	1.02924	0.90735	1.03984	1.47843
σ_Y^2	18.32858	9.43786	62.14178	-0.60189
σ_R^2	0.65276	0.36337	0.76928	1.30337
σ_π^2	0.44451	2.33699	3.60734	-7.64773
σ_C^2	10.3888	7.39139	63.64939	-1.00088
σ_I^2	71.79914	60.45211	284.8093	-1.12994
Wald stat	100	Mah. Distance (Normalised) 4.7		

Table 5: VAR Parameters, data variances and Model Bootstrap Bounds of the NC Model with Estimated Coefficients

in inflation and interest rates only a limited variance in demand shocks is required; and to mimic their persistence, the supply shocks must be sufficiently autocorrelated.

We have seen, however, that both the NK and NC versions of the SW model fail to match the data. Essentially, the NK model generates too little nominal variance while the NC model delivers too much. Given that each model fails in an opposite way, we propose a hybrid model that merges the NK and NC models by assuming that wage and price setters find themselves supplying labour and intermediate output partly in a competitive market with price/wage flexibility, and partly in a market with imperfect competition. We assume that the size of each sector depends on the facts of competition and do not vary in our sample but we allow the degree of imperfect competition to differ between labour and product markets.¹ We also assume

¹Formally, we model this as follows. We assume that firms producing intermediate goods have a production function that combines in a fixed proportion labour in imperfect competition ('unionised') with labour from competitive markets- thus the labour used by intermediate firms becomes $n_t = n_{1t} + n_{2t} =$

$$\left\{ \left[\int_0^1 (n_{1it})^{\frac{1}{1+\lambda_{w,t}}} di \right]^{1+\lambda_{w,t}} + \left[\int_0^1 (n_{2it}) di \right] \right\}$$

where n_{1it} is the unionised, n_{2it} the competitive labour provided by the i th household at t ; we can think of n_t as representing the activities of an intermediary 'labour bundler'. Note that $n_{1t} = v_w n_t$, $n_{2t} = (1 - v_w)n_t$ so that $W_t = v_w W_{1t} + (1 - v_w)W_{2t}$. Each household's utility includes the two sorts of labour in the same way, that is $U_{it} = \dots - \frac{n_{1it}^{1+\sigma_n} \epsilon_{1nt}}{1+\sigma_n} - \frac{n_{2it}^{1+\sigma_n} \epsilon_{2nt}}{1+\sigma_n} \dots$. W_{1t} is now set according to the Calvo wage-setting equation, while W_{2t} is set equal to current expected marginal monetary disutility of work; in the latter case a 1-quarter information lag is assumed for

that the monetary authority pursues a Taylor Rule that reflects the properties of the hybrid model.

In the hybrid model the price and wage setting equations are assumed to be a weighted average of the corresponding NK and NC equations. This weighting process is an informal use of indirect inference, the idea being to find the combination of the weights and Taylor coefficients that make the combined model perform best when compared with the auxiliary model.

We find that the optimal weights are $v_w = 0.1$ (the NK share for wages) and $v_p = 0.2$ (the NK share for prices). That is, only 10% of labour markets and only 20% of product markets are imperfectly competitive. Therefore, the model requires only a small amount of nominal rigidity in order to match the data. The Taylor rule then becomes: $R_t = 0.6R_{t-1} + (1 - 0.6)\{2.3\pi_t + 0.08y_t\} + 0.22(y_t - y_{t-1}) + \varepsilon_t$. This is a somewhat more aggressive response to inflation than either the NK ($R_t = 0.81R_{t-1} + (1 - 0.81)\{2.04\pi_t + 0.08(y_t - y_t^P)\} + 0.22[(y_t - y_t^P) - (y_{t-1} - y_{t-1}^P)] + \varepsilon_t^r$) or NC rules (the NC is the same as NK except that it sets ‘potential output’ to a constant). Notice that if one substitutes out for the interest rate from a simple money demand function with an exogenous money supply growth process, then one obtains a ‘Taylor Rule’ that has the form $\Delta R_t = \frac{1}{\beta}\{\pi_t + \gamma\Delta y_t\} + v_t$ where β is the semi-log interest rate elasticity of money demand, (γ is the corresponding income elasticity) and v_t is a combination of the money supply growth process and the change in the money demand error. This is fairly close to the rules adopted in these models when the lagged term in interest rates is large and the term in the output gap is small compared with the term in the rate of change of output.

The main difference between the hybrid and the NK and NC models is the hybrid model’s ability to reproduce the variances in the data. Using the structural errors from the model and the observed data, we find that all the data variances lie within the model’s 95% bounds (Table 6, last 5 entries). Furthermore, only nine of the 25 VAR coefficients lie outside their 95% confidence intervals. While the Full Wald statistic of 100 rejects this model version as it does the others, the Mahalanobis Distance of 3.1 implies that the hybrid model is substantially closer to the data.

Since the optimal combination indicates that the majority of the market participants behave in a competitive manner, it is not a surprise that the variance decomposition (Table 7) shows that the supply shocks - (productivity and labour supply shocks - explain most of the movements of the real variables. They also explain a large part of the nominal variables. While the demand shocks also contribute quite a lot to movements in the interest rate, they do so less for movements in inflation. So why are these results different from those of the NK and NC models?

The hybrid model mostly acts like the NC model, where the supply shocks explain most of the variation and demand shocks play a small part in the variability of real variables due to one period information lag and they add to the variability of nominal variables. Since, however, some economic agents behave in the New Keynesian manner, aggregate supply and labour supply are more elastic, demand shocks have a greater impact on real variables. Most importantly, inflation variability is dampened down to encompass actual data variability. It is remarkable how large the reduction in the lower bound is by the introduction of only small Calvo shares (10% in wages, 20% in prices — or 30% rigidity overall); the lower bound of inflation’s standard deviation falls no less than 57%. The reason appears to be that the variability of inflation also reacts to the variability of expected inflation. Thus, as the Calvo element rises, expected inflation varies less which, in turn, reduces the variability of actual inflation and, again in turn, reduces the variability of expected inflation, and so on in a sort of ‘multiplier’ process. This is an effect anticipated by Dixon (1992,1994).

Now we investigate the VAR impulse response functions to three main shocks: investment, labour supply,

current inflation but for convenience this is ignored in the usual way as unimportant in the Calvo setting over the whole future horizon.

These wages are then passed to the labour bundler who offers a labour unit as above at this weighted average wage. Firms then buy these labour units off the manager for use in the firm.

Similarly, retail output is now made up in a fixed proportion of intermediate goods in an imperfectly competitive market and intermediate goods sold competitively. Retail output is therefore $y_t = y_{1t} + y_{2t} =$

$\left\{ \left[\int_0^1 y_{j1t}^{\frac{1}{1+\lambda_{p,t}}} dj \right]^{1+\lambda_{p,t}} + \left[\int_0^1 y_{j2t} dj \right] \right\}$. The intermediary firm prices y_{1t} according to the Calvo mark-up equation on marginal costs, and y_{2t} at marginal costs.

Note that $y_{1t} = v_p y_t$, $y_{2t} = (1 - v_p)y_t$ so that $P_t = v_p P_{1t} + (1 - v_p)P_{2t}$. The retailer combines these goods as above in a bundle which it sells at this weighted average price.

Notice that apart from these equations the first-order conditions of households and firms will be unaffected by what markets they are operating in.

VAR coeffs	Actual Estimate	Lower Bound	Upper Bound	T-stats
A_Y^Y	0.99908	0.76761	0.99753	1.58317
A_Y^R	0.01503	-0.04394	0.01945	1.55148
A_Y^π	-0.00417	-0.02898	0.06556	-0.76799
A_Y^C	0.10174	-0.03905	0.10284	2.05348
A_Y^I	0.22591	-0.2093	0.28019	1.59406
A_R^Y	-0.64529	-0.97523	-0.14419	-0.40543
A_R^R	0.85001	0.49302	0.75838	3.23533
A_R^π	0.15154	-0.27545	0.1136	2.20900
A_R^C	-0.5553	-0.69948	-0.10079	-1.14167
A_R^I	-1.7064	-1.69603	0.44879	-1.94286
A_π^Y	0.11612	-0.29996	0.30247	0.72094
A_π^R	0.02374	0.07486	0.27241	-2.85460
A_π^π	0.59496	0.51488	0.78029	-0.76721
A_π^C	-0.38833	-0.27412	0.18435	-2.98726
A_π^I	-0.25917	-1.31646	0.30961	0.54641
A_C^Y	-0.08009	-0.142	0.09477	-0.82072
A_C^R	-0.02553	-0.06465	0.00228	0.19880
A_C^π	0.0121	-0.09455	0.00559	2.01856
A_C^C	0.78488	0.81382	0.96991	-2.92486
A_C^I	-0.4296	-0.36452	0.1334	-2.63914
A_I^Y	0.02034	-0.00329	0.07082	-0.56989
A_I^R	0.01022	0.00293	0.02417	-0.45314
A_I^π	0.01159	-0.01199	0.02095	0.83386
A_I^C	0.01957	-0.00749	0.03827	0.27887
A_I^I	1.02924	0.89898	1.04153	1.49582
σ_Y^2	18.32858	9.69749	61.85333	-0.61346
σ_R^2	0.65276	0.29191	0.76451	1.58414
σ_π^2	0.44451	0.43895	0.89102	-1.65685
σ_C^2	10.3888	7.30487	72.01693	-0.99793
σ_I^2	71.79914	61.41478	301.772	-1.17817
Wald stat	100	Mah. Distance (Normalised) 2.8		

Table 6: VAR Parameters, data variances and Model Bootstrap Bounds of the Weighted Model with Estimated Coefficients

Shocks	Govt. Spending	Pref	Inv	Mon	Prod	Price mark-up	Wage mark-up	Labour supply	Total
Y	2.6796	0.9823	1.9547	0.6995	48.2598	0.5086	0.00003	44.9154	100
R	11.8312	16.2245	17.4343	2.2156	15.3872	3.5352	0.000695	33.3713	100
π	2.0282	7.0541	3.7657	33.3303	17.6394	4.9596	0.000769	31.2218	100
C	5.0587	1.0009	1.7749	0.6429	34.3637	0.34915	0.00004	56.8097	100
I	11.298	0.0508	28.0270	0.1305	32.6701	0.2853	0.00001	27.5383	100

Table 7: Variance Decompositions of the weighted Model with estimated rhos

and productivity shocks. The main differences from the data are in the long-run responses of interest and inflation rates to the shocks; also the response of consumption is much more aggressive in the data than in the model. Nonetheless, these responses lie only just outside the 95% bounds. We can therefore say that the performance of the hybrid model, based on the IRFs, is relatively good compared to the NK and NC models.

The cross-correlations are accepted in a number of cases. The actual autocorrelations and cross-correlations of the variables lie within the model's bounds, though the correlation of investment with future output lies outside the bound. The performance of the cross-correlations among the nominal variables is, however, poor. The autocorrelations of interest and inflation rates are underpredicted by the model, even though the differences are much smaller than those for the NK and NC models. These failures are consistent with the overall rejection of the hybrid model.

We now consider the model's performance for particular aspects of the data, using the Directed Wald test. Our method is to focus first on individual variables and then in groups by estimating the best ARMA(i,j) in the case of a single variable and a VAR(1) for a group of variables. We then apply the Directed Wald test. To assess the individual shocks we take the IRFs (we use the IRF average) of the shock for the variables where they have a major impact and generate the model-implied joint distribution of these IRFs, computing the Wald statistic for the joint values in the data. We also look at the joint distribution of the variances to confirm our earlier judgement from the individual variances.

Table 8 below reports these Wald statistics. First, the model does fit the data variances jointly but only at the 99% level. Second, the real variables fit the data taken as a group, though again only at the 99% level, as do nominal variables taken as a group. When, however, nominal and real variables are combined the dynamic fit deteriorates sharply and the model is rejected at the 99% level; only if we restrict ourselves to output and inflation does the model pass this Wald test at the 99% level. This is mirrored in the individual shocks; the responses to both productivity and labour supply, the two key shocks in this model, are borderline rejected at 99%.

For individual variables, the responses of all are accepted at the 99% level; inflation is accepted at the 95% level. As observed earlier, many of the VAR coefficients involving interest rates are rejected individually. It therefore seems clear that this is the area to look for better specification of the model.

Variable combinations	Direct Wald
Y, C, INV	98.3
Y, C, INV, R	99.0
Y, C, INV, π	100
Y, R, π	99.4
Y, π	97.6
R, π	96.2
$Y (AR(3))$	96.2
$R (ARMA(1,1))$	98.4
$\pi (AR(3))$	90.3
$C (AR(3))$	98.8
$INV (AR(2))$	95.2

Table 8: Directed Wald statistics BY VARIABLE COMBINATIONS

Shocks	Variables	Directed Wald
$Prod$	Y, R, π, C, INV	98.2
$LabSup$	Y, R, π, C, INV	99.1
Variances		Directed Wald
$\sigma_Y^2, \sigma_R^2, \sigma_\pi^2, \sigma_C^2, \sigma_I^2$		97

Table 9: Directed Wald statistics

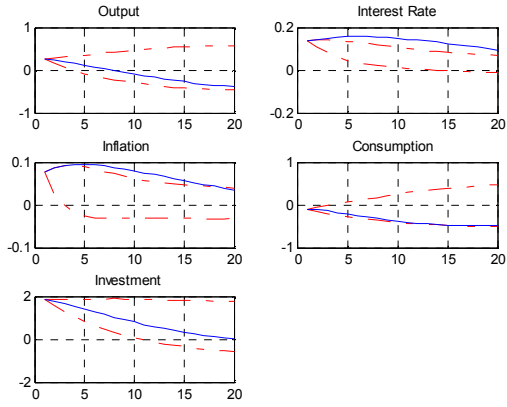


Figure 3: Investment Shock

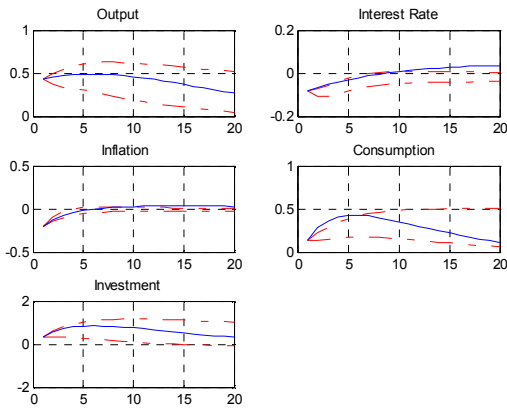


Figure 4: Productivity Shock

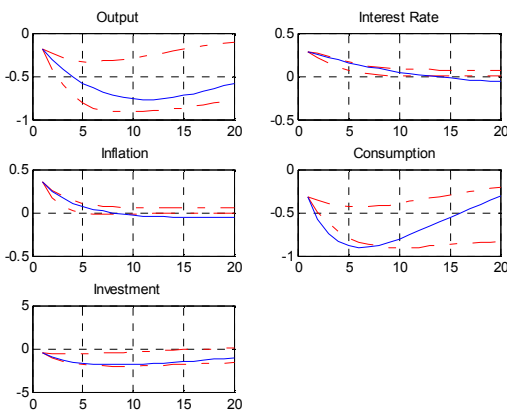


Figure 5: Labour Supply Shock

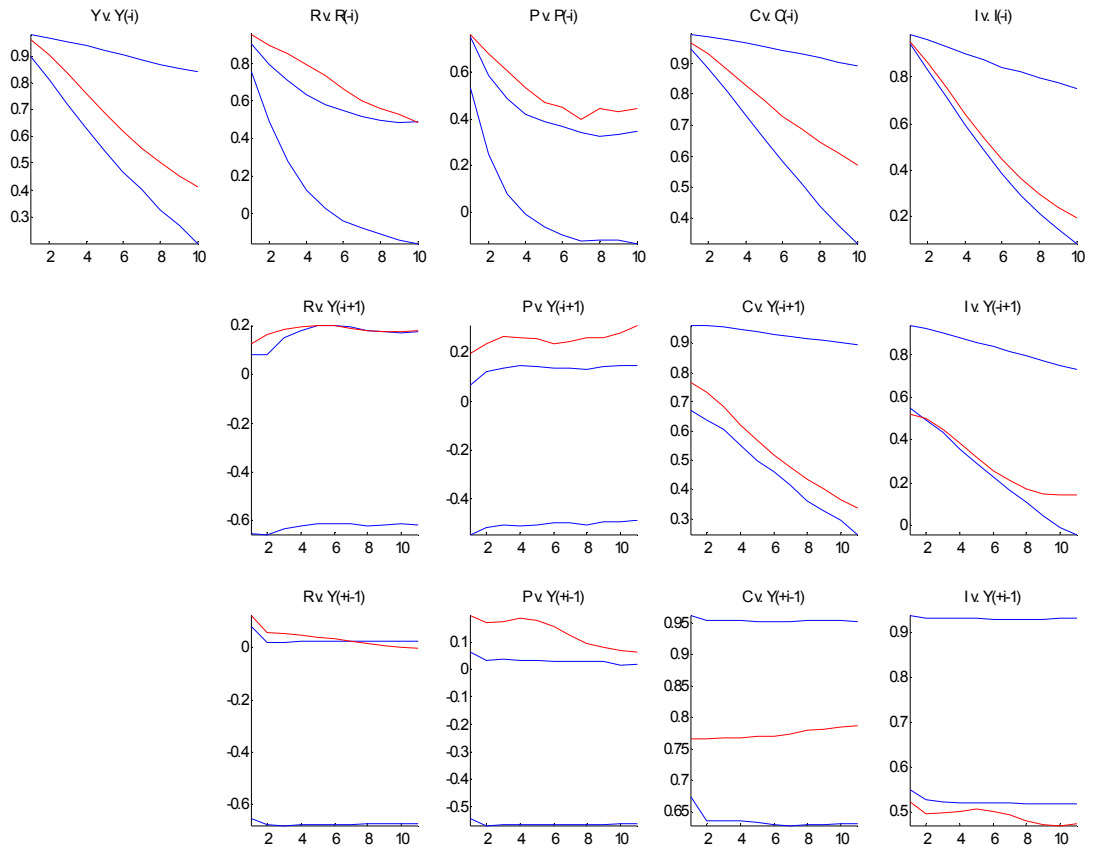


Figure 6: Cross-Correlations for Weighted Model (with estimated coefficients)

5 Regime change as a possible source of mis-specification

In view of the apparently crucial role of interest rates in the hybrid model, the implication is that the problem could lie in the specification of monetary policy, and in particular the use of one monetary regime for the whole sample from 1950s to the 2000s. We therefore tested for structural change during this period following the procedure of Perron and Wu (2007) designed to test for multiple breaks in VAR parameters; we found evidence of parameter breaks in two places: 1965 and 1984.

These are natural places to find such breaks due to changes that occurred in the monetary regime. The earlier break is associated with the emergence of serious inflation for the first time; the later break is associated with the shift towards interest rate setting that followed from the adoption of (implicit) inflation targeting.

In moving to three sub-periods we tripled the size of our testing problem. Furthermore linear detrending no longer proved sufficient to make the data stationary; we therefore used a Hodrick-Prescott filter. So far

The estimated breaks are:	1965.02	1984.02
The 95% C.I. for the 1st break is	(1964.04;1965.04)	
The 95% C.I. for the 2nd break is	(1983.02-1985.02)	

Table 10: Perron-Wu Multivariate Structural Break Test

	Actual	Lower	Upper	State
A_Y^Y	0.92801	0.61496	0.94141	<i>TRUE</i>
A_Y^R	0.06743	-0.04377	0.04308	<i>FALSE</i>
A_Y^π	0.05812	-0.03036	0.06507	<i>TRUE</i>
A_R^Y	-0.47031	-0.91141	0.53397	<i>TRUE</i>
A_R^R	0.75002	0.51648	0.87073	<i>TRUE</i>
A_R^π	0.09010	-0.06550	0.38117	<i>TRUE</i>
A_π^Y	0.13438	-1.06434	0.62982	<i>TRUE</i>
A_π^R	0.12020	-0.17840	0.29051	<i>TRUE</i>
A_π^π	0.00798	-0.12482	0.35663	<i>TRUE</i>
$Var(Y)$	0.91360	0.54450	2.08374	<i>TRUE</i>
$Var(R)$	0.06939	0.03334	0.11078	<i>TRUE</i>
$Var(\pi)$	0.03058	0.02884	0.05445	<i>TRUE</i>
Wald	96.8	M-distance (Normalised)		1.9

Table 11: VAR coefficients and variances

	Direct Walds
$Y (AR(3))$	54.7
$R (AR(1))$	97.2
$R (AR(2))$	100
$\pi (AR(1))$	69.6
Y, π	88.5
Y, R, π	96.6
$var(Y), var(R), var(\pi)$	60.5

Table 12: Direct Walds for different combinations of output, inflation and interest rate

we have been unable to locate acceptable versions of the model for the first two sub-periods. However for the third and latest sub-period (1984.03-2004.02), we found good results when we shifted the weights in the hybrid model greatly towards the New Keynesian end of the spectrum (0.8, 0.8). It may well be that in the 'great moderation' price-setting was far less disturbed by shocks to the state and was dominated instead by time dependence.

The model is still rejected on the full VAR with a Wald of 100 but its Mahalanobis Distance is the lowest to this point, at 4.2 (similar to the full sample weighted model of 3.9 on H-P-filtered data). While it is also rejected for interest rates alone for its best AR(2) representation, for the combined variable set of output, inflation and interest rates it is now jointly accepted at the 99% level and nearly accepted at the 95% level. For just output and inflation it is easily accepted at the 95% level (88.5). Significantly, this is the first time that any model we have examined over the full data period has passed the test embracing real GDP and both nominal variables.

6 Conclusion

We have used the method of indirect inference to test a well-known DSGE model of the New Keynesian type on its dynamic performance for US post-war data. We compared this model with a flexible wage/price version with a short information lag (New Classical) and found that if we use the structural errors jointly implied by each model and the data, then neither model can fit the data variances. The NK produces too little variation in interest rates and the NC model generates an excessive variation in inflation rates. But when the two models are combined in a weighted combination to give a hybrid model which is a mixture of imperfectly competitive and flexible-price markets, then with 90% flexibility in the product market and 80% in the labour market the hybrid model comes much closer to matching the data, even though it too is rejected as mis-specified especially in respect of interest rate behaviour. One possible reason is monetary regime change, for which there is evidence in the mid-1960s and the mid-1980s. When we examined the

period since 1984 we found that in respect of output and nominal variables the data did not reject a model with quite a high degree of nominal rigidity (around 90% in both goods and labour markets). This suggests that the situation with regard to price or wage rigidity in the US economy may have changed over time; during most of the period state dependence in pricing could have dominated time dependence due to the economy's large fluctuations, but, during the later period, the 'great moderation', time dependence seems to have dominated.

References

- [1] Bils, M. and P. J. Klenow, 2004, "Some Evidence on the Importance of Sticky Prices," *Journal of Political Economy*, 112 (5), 947–985.
- [2] Canova, F., 2005. *Methods for Applied Macroeconomic Research*, Princeton University Press, Princeton.
- [3] Canova, F. and L. Sala (2009) 'Back to square one: Identification issues in DSGE models', *Journal of Monetary Economics*, 56, 431–449.
- [4] Canzoneri, M. , R. Cumby, and B. Diba, 2004. "The Cost of Nominal Inertia in NNS Models," NBER Working Papers 10889, National Bureau of Economic Research, Inc
- [5] Christiano, L., M. Eichenbaum, and C. Evans, 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, University of Chicago Press, vol. 113(1), pages 1–45, February.
- [6] Del Negro, M., F. Schorfheide, F. Smets, R. Wouters (2006) 'On the fit of new Keynesian models', *Journal of Business and Economic Statistics*, 25, 143–162.
- [7] Dixon, H. 1992. "Nominal wage flexibility in a partly unionised economy," *The Manchester School of Economic and Social Studies*, 1992, 60, 295–306.
- [8] Dixon, H. 1994. "Macroeconomic price and quantity responses with heterogeneous product markets," *Oxford Economic Papers*, 1994, 46, 385–402.
- [9] Gertler, M. and J. Leahy (2008) "A Phillips curve with an Ss foundation", *Journal of Political Economy*, 118 (3), 533–572.
- [10] Gourieroux, C. and A., Monfort, 1995. *Simulation Based Econometric Methods*, CORE Lectures Series, Louvain-la-Neuve.
- [11] Gourieroux, C., A., Monfort, and E., Renault, 1993. "Indirect inference", *Journal of Applied Econometrics*, 8, 85–118.
- [12] Gregory, A. and G. Smith, 1991. " Calibration as testing: Inference in simulated macro models," *Journal of Business and Economic Statistics*, 9, 293–303.
- [13] Gregory, A. and G. Smith, 1993. "Calibration in macroeconomics", in G. Maddala, ed., " *Handbook of Statistics*", vol. 11, Elsevier, St. Louis, Mo., 703–719.
- [14] Juillard, M. (2001) "Dynare: a program for the simulation of rational expectations models" *Computing in economics and finance* 213. Society for Computational Economics.
- [15] Kimball, M., 1995. "The quantitative analytics of the basic neomonetarist model", *Journal of Money, Credit and Banking*, 27(4), Part 2, 1241–1277.
- [16] Le, V. P. M., 2008. "Dynamic Races between New Keynesian and New Classical models of the US," Cardiff University, Mimeo.
- [17] McCallum, B. T. (1976) "Rational expectations and the natural rate hypothesis: some consistent estimates", *Econometrica*, 44, 43–52.

- [18] Meenagh, D. , P. Minford, and M. Wickens, 2009, "Testing a DSGE model of the EU using indirect inference," *Open Economies Review*, 20(4), 435-471.
- [19] Perron, P. and Qu, Z. (2007) "Estimating and Testing Structural Changes in Multivariate Regressions," *Econometrica*, 75(2), 459-502
- [20] Smets, F. & R. Wouters, 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, MIT Press, vol. 1(5), pages 1123–1175, 09.
- [21] Smets, F. & R. Wouters, 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, American Economic Association, vol. 97(3), pages 586–606, June
- [22] Smith, A. , 1993. " Estimating nonlinear time-series models using simulated vector autoregressions", *Journal of Applied Econometrics*, 8, S63–S84.
- [23] Wickens, M.R. (1982) "The efficient estimation of econometric models with rational expectations", *Review of Economic Studies*, 49, 55-67.

Appendix A Listing of models — SWNK and SWNC

SWNK MODEL

Consumption Euler equation

$$c_t = \frac{\frac{0.71}{1.0043}}{1 + \frac{0.71}{1.0043}} c_{t-1} + \frac{1}{1 + \frac{0.71}{1.0043}} E_t c_{t+1} + \frac{(1.3952 - 1) 0.83192}{\left(1 + \frac{0.71}{1.0043}\right) 1.38} (l_t - E_t l_{t+1}) - \left(\frac{1 - \frac{0.71}{1.0043}}{\left(1 + \frac{0.71}{1.0043}\right) 1.38}\right) (r_t - E_t p_{t+1}) + e b_t \quad (1)$$

Investment Euler equation

$$inn_t = \frac{1}{1 + 0.99(1.0043)} inn_{t-1} + \frac{0.99(1.0043)}{1 + 0.99(1.0043)} E_t inn_{t+1} + \frac{1}{(1 + 0.99(1.0043)) (1.0043)^2 5.74} qq_t + e inn_t \quad (2)$$

Tobin Q equation

$$qq_t = \frac{1 - 0.025}{1 - 0.025 + 0.032649} E_t qq_{t+1} + \frac{0.032649}{1 - 0.025 + 0.032649} E_t r k_{t+1} - (r_t - E_t p_{t+1}) + \frac{1}{\frac{1 - \frac{0.71}{1.0043}}{\left(1 + \frac{0.71}{1.0043}\right) 1.38}} e b_t \quad (3)$$

Capital Accumulation equation

$$k_t = \left(1 - \frac{0.025}{1.0043}\right) k_{t-1} + \frac{0.025}{1.0043} inn_t + \left(1 - \frac{0.025}{1.0043}\right) (1 + 1.0043 (0.99)) (1.0043)^2 (5.74) (e nn_t) \quad (4)$$

Price Setting equation

$$p_t = \left[\frac{\frac{0.99x1.0043}{1+0.99x1.0043 \times 0.24} E_t p_{t+1} + \frac{0.24}{1+0.99x1.0043 \times 0.24} p_{t-1} - \left(\frac{1}{1+0.99x1.0043 \times 0.24}\right)}{\left(\frac{(1-0.99(1.0043)(0.66))(1-0.66)}{0.66((1.6-1)(10)+1)}\right)} (0.19 r k_t + (1 - 0.19) w_t - e a_t) \right] + e p_t \quad (5)$$

Wage Setting equation

$$w_t = \left[\frac{\frac{0.99x1.0043}{1+0.99x1.0043} E_t w_{t+1} + \frac{1}{1+0.99x1.0043} w_{t-1} + \frac{0.99x1.0043}{1+0.99x1.0043} E_t p_{t+1} - \frac{1+0.99x1.0043 \times 0.58}{1+0.99x1.0043} p_t}{\frac{0.58}{1+0.99x1.0043} p_{t-1} - \frac{1}{1+0.99x1.0043} \left(\frac{(1-0.99x1.0043 \times 0.7)(1-0.7)}{(1+(10)(1.5-1))0.7}\right)} + \frac{1}{1+0.99x1.0043} \left(\frac{1-0.71}{1.0043}\right) (c_t - \frac{0.71}{1.0043} c_{t-1})} \right] + e w_t \quad (6)$$

Labour demand

$$l_t = -w_t + \left(1 + \frac{1 - 0.54}{0.54}\right) r k_t + k_{t-1} \quad (7)$$

Market Clearing condition in goods market

$$y_t = 0.64 c_t + 0.17 inn_t + 0.19 \frac{1 - 0.54}{0.54} r k_t + e g_t \quad (8)$$

Aggregate Production equation

$$r k_t = \frac{1}{1.6 (0.19) \frac{1-0.54}{0.54}} (y_t - 1.6 (0.19) k_{t-1} - 1.6 (1 - 0.19) l_t - 1.6 e a_t) \quad (9)$$

Taylor Rule

$$r_t = 0.81 r_{t-1} + (1 - 0.81) \left(2.03 p_t + 0.08 (y_t - y_t^f)\right) + 0.22 \left(\left(y_t - y_t^f\right) - \left(y_{t-1} - y_{t-1}^f\right)\right) + e r_t \quad (10)$$

SWNC MODEL

Consumption Euler equation

$$c_t = \frac{\frac{0.71}{1.0043}}{1 + \frac{0.71}{1.0043}} c_{t-1} + \frac{1}{1 + \frac{0.71}{1.0043}} E_t c_{t+1} + \frac{(1.3952 - 1) 0.83192}{(1 + \frac{0.71}{1.0043}) 1.3952} (l_t - E_t l_{t+1}) - \left(\frac{1 - \frac{0.71}{1.0043}}{(1 + \frac{0.71}{1.0043}) 1.3952} \right) (r_t) + e b_t \quad (11)$$

Investment Euler equation

$$inn_t = \frac{1}{1 + 0.99(1.0043)} inn_{t-1} + \frac{0.99(1.0043)}{1 + 0.99(1.0043)} E_t inn_{t+1} + \frac{1}{(1 + 0.99(1.0043)) (1.0043)^2 5.74} qq_t + e inn_t \quad (12)$$

Tobin Q equation

$$qq_t = \frac{1 - 0.025}{1 - 0.025 + 0.032649} E_t qq_{t+1} + \frac{0.032649}{1 - 0.025 + 0.032649} E_t r k_{t+1} - (r_t) + \frac{1}{\frac{1 - \frac{0.71}{1.0043}}{(1 + \frac{0.71}{1.0043}) 1.3952}} e b_t \quad (13)$$

Capital accumulation equation

$$k_t = \left(1 - \frac{0.025}{1.0043} \right) k_{t-1} + \frac{0.025}{1.0043} inn_t + \left(1 - \frac{0.025}{1.0043} \right) (1 + 1.0043 (0.99)) (1.0043)^2 (5.74) (enn_t) \quad (14)$$

Marginal Product of Labour

$$0.19 r k_t + (1 - 0.19) w_t = e a_t \quad (15)$$

Labour supply

$$w_t = 1.83 l_t + \left(\frac{1}{1 - \frac{0.71}{1.0043}} \right) (c_t - \frac{0.71}{1.0043} c_{t-1}) - (\pi_t - E_{t-1} \pi_t) \quad (16)$$

Labour Demand

$$l_t = -w_t + \left(1 + \frac{1 - 0.54}{0.54} \right) r k_t + k_{t-1} \quad (17)$$

Market clearing condition

$$y_t = 0.64 c_t + 0.17 inn_t + 0.19 \frac{1 - 0.54}{0.54} r k_t + e g_t \quad (18)$$

Production function

$$r k_t = \frac{1}{1.6 (0.19) \frac{1 - 0.54}{0.54}} (y_t - 1.6 (0.19) k_{t-1} - 1.6 (1 - 0.19) l_t - 1.6 e a_t) \quad (19)$$

Taylor Rule

$$r_t = 0.81 r_{t-1} + (1 - 0.81) (2.03 p_t + 0.08 y_t) + 0.22 (y_t - y_{t-1}) + e r_t \quad (20)$$