**Cardiff Economics
Working Papers**

David Meenagh, Patrick Minford and Michael Wickens

Testing a DSGE Model of the EU Using Indirect Inference

E2008/11

Testing a DSGE model of the EU using indirect inference[*]

David Meenagh

Cardiff University

Patrick Minford[†]

Cardiff University and CEPR

Michael Wickens

Cardiff University, University of York and CEPR

December 2008

Abstract

We use the method of indirect inference, using the bootstrap, to test the Smets and Wouters model of the EU against a VAR auxiliary equation describing their data. We find that their model generates excessive variance compared with the data. But their model fits the dynamic facts quite well if the errors have the properties assumed by SW but scaled down. We compare a New Classical version of the model which also performs reasonably if error properties are chosen using New Classical priors (notably excluding shocks to preferences). Both versions have (different) difficulties fitting the data if the actual error properties are used. A model combining rigid and flexible-wage/price sectors, with a weight of around 5% on the rigid sector, does best in fitting the data.

JEL Classification: *C12, C32*

Keywords: *Bootstrap, DSGE Model, VAR model, Model of EU, indirect inference, Wald statistic*

1   Introduction

In a notable recent contribution Smets and Wouters (2003) proposed a dynamic stochastic general equilibrium (DSGE) model of the EU which they estimated by Bayesian methods after allowing for a complete set of pre-specified, but ad hoc, stochastic shocks. They reported that, based on measures of fit and dynamic performance, their model was superior in performance both to a Bayesian and a standard VAR. In this paper we look carefully at their innovative model and review its performance, using a new evaluation procedure that is suitable for either a calibrated or, as here, an estimated structural model. The method is based on indirect inference. It exploits the properties of the model's error processes through bootstrap simulations. We ask whether the simulated data of a calibrated or an estimated structural model, treated as the null hypothesis, can explain the actual data where both are represented by the dynamic behaviour of a well-fitting auxiliary model such as a VAR. Our proposed test statistic is a multi-parameter portmanteau Wald test that focuses on the structural model's overall capacity to replicate the data's dynamic performance.

The Smets-Wouters (SW) model follows the model of Christiano et al. (2005) for the US but is fitted to the data using Bayesian estimation methods that allow for a full set of shocks. It is a New-Keynesian model, i.e. it is based on the New Neo-Keynesian Synthesis involving a basic Real Business Cycle framework under imperfect competition in which there are menu costs of price and wage change modelled by Calvo contracts and a backward-looking indexation mechanism; monetary policy is supplied by an interest-rate setting rule. The effect is to impart a high degree of nominal rigidity to the model, both of prices and inflation. A central tenet of New-Keynesian authors is that this is necessary in order to fit the dynamic properties of the data which are characterised by substantial persistence in output and inflation, and hump-shaped responses to monetary policy shocks. In this paper we probe this argument. Specifically, we compare the SW model with a flexprice version in which prices and wages are flexible and there is a physical one quarter lag in the arrival of macro information. Thus our alternative model is a type of 'New Classical' model (SWNC). We also assess the contribution to the success of their structural model of the ad hoc structural shocks assumed by Smets and Wouters.

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 model. In estimation the idea is to choose the parameters of the structural model 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 given. The aim is to compare the performance of the auxiliary model estimated on simulated data from the given structural model with the performance of the auxiliary model when estimated from 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.

We find that the properties of the prior distributions of the parameters and the stochastic shocks, whether assumed or generated from the observed data, are the key element in the success or failure of both the Smets-Wouters model and our New Classical variant. The more the error properties conform to New Keynesian priors where there is considerable price stickiness and demand shocks are dominant, the better the Smets-Wouters model performs and the worse our New Classical is. But, in contrast, the more the errors conform to New Classical priors, the better this model performs and the worse the Smets-Wouters model is. Further investigation reveals that only a small degree of price stickiness incorporated in the New Classical model is sufficient to allow it to fit the data well. This suggests that price stickiness, whilst relevant, is not as important as might be suggested by the results of Smets and Wouters.

We begin by describing our model evaluation procedure in section 2 and relate it to estimation by indirect inference. In section 3 we describe the SW model and its findings based on its own detrended data (which we use throughout this paper). In section 4 we apply our proposed testing procedure to the model and compare its performance with a New Classical version of the model. Section 5 compares the New Keynesian and New Classical models and introduces a weighted combination of the two, and section 6 concludes.

2 Model evaluation by indirect inference

The 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 the 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, we discuss estimation by indirect inference.

2.1 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, $t = 1, ..., T$, $x_t(\theta)$ 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.

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

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

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

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

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

with

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

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

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

4

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

$$plim\ a_T = plim\ a_S(\theta) = \alpha$$

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

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

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

The extended method of simulated moments estimator (EMSME) 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 $G_T(a_T) = \frac{1}{T}\Sigma_{t=1}^T g(a_T)$ and $G_S(\alpha_S(\theta)) = \frac{1}{S}\Sigma_{s=1}^S g(\alpha_S(\theta))$. 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

$$\theta_{T,S} = \arg\min_{\theta}[G_T(a_T) - G_S(\alpha_S(\theta))]'W(\theta)[G(a_T) - G_S(\alpha_S(\theta))]$$

2.2 Model evaluation

The parameters of the macroeconomic model and their distributions are now 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 from the given distributions of the parameters. The test statistic is based on the distributions of these functions of the parameters of the auxiliary model, or of a function of these parameters. We choose the auxiliary model to be a VAR and base our test on a function of the VAR coefficients.

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.

A 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 $\widehat{\theta}$ is the estimated or calibrated value of θ then the null hypothesis can be expressed as $H_0 : \widehat{\theta} \rightarrow$

θ_0. Consider again the continuous $p \times 1$ vector of functions $g(a_T)$, $g(\alpha_S(\theta))$, $G_T(a_T) = \frac{1}{T}\Sigma_{t=1}^{T}g(a_T)$ and $G_S(\alpha_S(\theta)) = \frac{1}{S}\Sigma_{s=1}^{S}g(\alpha_S(\theta))$. The functions $g(.)$ 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 $G_T(a_T) - G_S(\alpha_S(\widehat{\theta}))$. The resulting Wald statistic is

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

where the estimate of the optimal weighting matrix is

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

Alternatively, the distribution of $G_T(a_T) - G_S(\alpha_S(\widehat{\theta}))$ and the Wald statistic can be obtained using the bootstrap. We take the following steps in our implementation of the Wald test by bootstrapping:

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

Solve the DSGE macroeconomic model for the structural the errors ε_t given $\widehat{\theta}$ 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 Normal

Step 2: Construct 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. This is the case with the SW model; we discuss below the precise assumptions made.

Step 3: Compute the Wald statistic

The test is here based on a comparison of the VAR coefficient vector itself rather than a multi-valued function of it such as the IRFs. Thus

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

also therefore

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

The distribution of $a_T - \alpha_S(\widehat{\theta})$ and its covariance matrix $W(\widehat{\theta})^{-1}$ are estimated by bootstrapping $\alpha_S(\widehat{\theta})$. This proceeds by drawing N bootstrap samples of the structural model, and estimating the auxiliary VAR on each, thus obtaining N $a_S(\widehat{\theta})$. This set of vectors represents the sampling

6

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 $[a_T - \alpha_S(\widehat{\theta})]'W(\widehat{\theta})[a_T - \alpha_S(\widehat{\theta})]$. We also compute this statistic with a diagonalised covariance matrix; this amounts to the sum of the squared t-statistics of the a coefficients, taken individually, and can be treated as a measure of the closeness of the estimated coefficients to the bootstrap means — which we call a 'joint t-test' (see below for more details.) It turns out in what follows that virtually all the model versions we examine fail the Wald test, while varying with respect to the joint t-test; hence the latter is useful as an overall ranking criterion.

Because the Smets-Wouters model is a log-linearised rational expectations DSGE model, and hence there is an exact VAR representation of the model, in principle, analytic methods could be used rather than the bootstrap. Nonetheless, we prefer the bootstrap because we wish to preserve the actual residuals implied by the structural model. We also draw the shocks by time vector, so as to preserve their contemporaneous relation.

We use a VAR(1) as our principal auxiliary model to provide a parsimonious description of the 'dynamic facts' against which several models may be compared with high power of discrimination for their ability to match this description. We also look at higher order VARs as supplementary information but we find they add nothing to our assessments.

2.2.1   The method in visual form:

To illustrate the bootstrap evaluation method and the role of the covariance matrix, we show below a three-dimensional figure for the parameter distribution of a VAR with just two parameters, for example inflation and interest rates regressed only on their own individual past (a diagonalised VAR). Suppose that the model distribution is centred around 0.5, and 0.5; and the data-based VAR produced values for their partial autocorrelations of 0.1 and 0.9 respectively for inflation and interest rates — the two VAR coefficients. Suppose too that the 95% range for each was 0–1.0 (a standard deviation of 0.25) and thus each is accepted individually. If the parameters are uncorrelated across samples, then the situation is as illustrated in the 3-D diagram below. The height of the diagram shows the density of parameter combinations across the samples. Here the mean of each parameter's distribution remains constant regardless of the value of the other parameter. Of course the joint parameter combination will also be accepted because of this independence.

Now consider the case where there is a high positive correlation between the parameter estimates across samples. Thus suppose that in samples with high inflation autocorrelation we also find high

interest rate autocorrelation (because of the Fisher effect perhaps). For example Figure 1 illustrates the case for a 0.9 correlation between the two parameters. The effect of the high correlation is to create a 'ridge' out of the 'density mountain'. Hence at high values of the interest rate autocorrelation the mean of the inflation autocorrelation is now increased from 0.5; for example at an interest rate parameter of 0.9 the mean of the inflation parameter distribution will be 0.86; the distance of 0.1 from a mean of 0.86 is 3.04 standard deviations. Thus the joint parameter combination of 0.1,0.9 will be rejected even though individually the two parameters are accepted.

When there are numerous VAR parameters, each pair will have the characteristics just described, creating ridges in multiple dimensions. The joint distribution of the VAR parameters will clearly in general have many such ridges, reflecting the way in which the parameters are jointly estimated in the various bootstrap samples in order to fit the bootstrap sample data. For example, here if we ignore the covariances (diagonalise the covariance matrix), we find that the a large proportion of the implied joint parameters in the distribution generate instability.
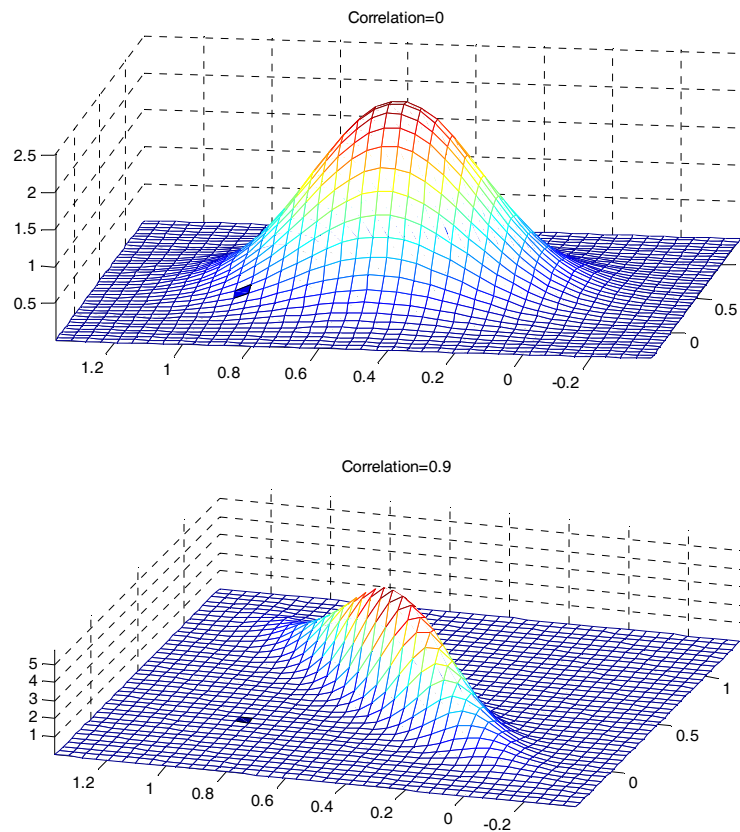

Figure 1: Bivariate Normal Distributions (0.1, 0.9 shaded) with correlation of 0 and 0.9.

3 The Smets-Wouters DSGE model of the EU

Following a recent series of papers Smets and Wouters (2003), SW have developed a DSGE model of the EU. This is in most ways an RBC model but with additional characteristics that make it 'New Keynesian'. First there are Calvo wage- and price-setting contracts under imperfect competition in labour and product markets, together with lagged indexation. Second, there is an interest-rate setting rule with an inflation target to set inflation. Third, there is habit formation in consumption. The model is described in full in the Appendix A.

Ten exogenous shocks are added to the model. Eight — technical progress, preferences and cost-push shocks — are assumed to follow independent AR(1) processes. The whole model is then estimated using Bayesian procedures on quarterly data for the period 1970q1–1999q2 for seven euro-area macroeconomic variables: GDP, consumption, investment, employment, the GDP deflator, real wages and the nominal interest rate. It is assumed that capital and the rental rate of capital are not observed. By using Bayesian methods it is possible to combine key calibrated parameters with sample information. Rather than evaluate the DSGE model based only on its sample moment statistics, impulse response functions are also used. The moments and the impulse response functions for the estimated DSGE model are based on the median of ten thousand simulations of the estimated model. A third-order VAR is fitted to the original data and is used to provide the impulse response functions for the original data. We now summarise the main findings of Smets and Wouters.

Comparing the auto-covariances of the VAR and the simulated DSGE model, those from the VAR are generally quite close to those of the DSGE model. The VAR auto-covariances lie within the confidence bands of those for the DSGE model; the bands are, however, quite wide, indicating parameter uncertainty. The main discrepancy concerns the auto-covariances between output and the expected real interest rate. These are higher in the VAR, but the differences are not significant.

Turning to the impulse response functions for the DSGE model, first we consider the responses to a positive productivity shock, ε_t^a. This causes output, consumption and investment to rise, but employment and the utilisation of capital to fall. The real wage also rises, but only gradually. The fall in employment is consistent with evidence on the impulse responses to US productivity shocks, but is in contrast to the predictions of the standard RBC model without nominal rigidities. A possible explanation is that, due to the rise in productivity, marginal cost falls on impact and, as monetary policy does not respond strongly enough to offset this fall, inflation declines gradually. The estimated reaction of monetary policy to a productivity shock is comparable to results for the US.

A positive labour supply shock has a similar effect on output, inflation and the interest rate to a positive productivity shock. Due to the higher persistence of the labour supply shock, the real interest rate is not greatly affected. The main differences compared to a standard RBC model are first that employment also rises in line with output and, second, that the real wage falls significantly. This fall in the real wage leads to a fall in marginal cost and in inflation. A negative wage mark-up shock has similar effects to these, except that the real interest rate rises, and real wages and marginal costs fall more on impact. The effects of a negative price mark-up shock on output, inflation and interest rates are also similar, but the effects on real marginal cost, real wages and the rental rate of capital are opposite in sign.

Positive demand shocks generally cause real interest rates to rise. A positive preference shock, while increasing consumption and output, crowds-out investment. The increase in capacity necessary to satisfy increased demand is delivered by an increase in the utilisation of installed capital and an increase in employment. Increased consumption demand puts pressure on the prices of the factors of production, and both the rental rate on capital and the real wage rise thereby putting upward pressure on marginal cost and inflation.

A positive government expenditure shock raises output initially, but crowds-out consumption which, due to increases in the marginal utility of working, leads to a greater willingness of households to work. As a result the effects on real wages, marginal costs and prices are small.

A negative monetary policy shock (increase in the interest rate shock η_t^R) has temporary effects on all variables apart from the price level, which falls permanently. For the first few periods, nominal and real short-term interest rates rise, and output, consumption, investment and real wages fall. The maximum effect on investment is about three times as large as that on consumption. Overall, these effects are consistent with other evidence on the euro area, though the price effects in the model are somewhat larger than those estimated in some identified VARs.

A permanent increase in target inflation ($\bar{\pi}_t$) does not have a strong effect on output, consumption, employment, the real wage or the real interest rate, although all rise quickly. It has a larger effect on investment and, of course, causes the price level to rise permanently.

The contribution of each of the structural shocks to variations in the endogenous variables may be obtained from the forecast error variances at various horizons. At the one-year horizon, output variations are driven primarily by the preference shock and the monetary policy shock. In the medium term, both of these shocks continue to dominate, but the two supply shocks (productivity and labour supply) account for about 20% of the forecast error variance. In the long run, the labour supply shock dominates, but the monetary policy shock still accounts for about a quarter of the forecast error in output. The monetary policy shock is transmitted mainly through investment.

The price and wage mark-up shocks make little contribution to output variability. Taken together, the two supply shocks, the productivity and the labour shock, account for only 37% of the long-run forecast error variance of output, which is less than is found in most VAR studies. The limited importance of productivity shocks, which explain a maximum of 12% of the forecast error variance of output, is probably due to the negative correlation between output and employment.

In the short run, variations in inflation are mainly driven by price mark-up shocks. This appears to be a very sluggish process, with inflation only gradually responding to current and expected changes in marginal cost. In the medium and long run, preference shocks and labour supply shocks account for about 20% of the variation in inflation, whereas monetary policy shocks account for about 15%.

In summary, in this study by Smets and Wouters three structural shocks explain a significant fraction of output, inflation and interest rates at the medium to long-term horizon: the preference shock, the labour supply shock and the monetary policy shock. In addition, the price mark-up shock is an important determinant of inflation, but not of output, while the productivity shock determines about 10% of output variations, but does not affect inflation. Smets and Wouters do not report corresponding results for government expenditure shocks, though these shocks appear to have a strong temporary effect on output. This suggests that RBC models, with their focus on productivity shocks, do not give an adequate representation of the economy, or even of output, and that the effects of monetary and, possibly, fiscal policy should also be represented in a DSGE macroeconomic model together with labour supply effects.

Comparing the estimated parameters with those obtained by other studies, their general conclusion is that they are similar. They also report various other tests: They compare the DSGE model with VAR(p) (where p=1,2,3) and Bayesian VAR(p) through the use of the Marginal Likelihood or the Bayes factor. They also provide impulse response and variance decomposition analysis. Their assessment is that the model behaves satisfactorily.

4 Testing the SW Model using the method of indirect inference

We now apply our proposed testing procedure to this model using throughout the same data for the period 1970–1999 as SW and the same detrended series obtained by taking deviations of all variables from a mean or a linear trend. We appear to replicate the solution of their model with reasonable accuracy; the method used is Dynare (Juillard (2001)). The distribution of SW's impulse response functions (IRFs) are obtained from repeated draws out of the structural parameter

posterior distributions. The IRFs of our version of their model have been produced from the median of the posterior distribution of their structural parameters, using the loglinearised version of their model that they too used. Of course we cannot know how the IRFs from any particular parameter combination will compare with their distribution; however the combination of median parameters might reasonably be expected to produce IRFs that for the most part lie inside the 95% bounds shown — and this seems to be the case (see Annex Alph@section for details).

We begin by estimating a VAR on the observed data, using the five main observable variables: inflation (quarterly rate), interest rate (rate per quarter), output, investment and consumption (capital stock, equity returns, and capacity utilisation are all constructed variables, using the model's identities; we omit real wages and employment from the VAR) all in units of percent deviation from trend. We focus on a VAR(1) in order to retain power for our tests, this yields 25 coefficients, apart from constants.

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

Our first evaluation of the SW model in its New Keynesian form (SWNK) uses the error properties they themselves assumed (i.e. obtained as the posteriors from their Bayesian estimation procedure). This is an important starting point as the properties they report for the model are based on these assumptions and not on the actual errors we will discuss shortly (we need to scale their errors by 0.25). Given that they have estimated the model satisfactorily with these assumed errors, we would expect it to perform well. This is indeed the case.

We notice first that the model's variance bounds comfortably encompass all the variables we focus on. We then find that the model passes the joint t-test with a statistic of 75.7; all the VAR coefficients bar one lie inside the 95% model bounds. With two marginal exceptions, the IRFs of the VAR (when identified by the model) also lie within the model bounds (again from the model bootstrap distribution of VAR coefficients). The two exceptions are the effect of a consumption shock on inflation and interest rates. Similarly, the gross cross-correlations in the data largely lie within the model bounds. The exceptions are, however, more serious. The interest rate autocorrelation structure and the correlation structure from output to interest rates are under-predicted, as is the correlation structure between consumption and output. Though most of the discrepancies are modest (less than 0.2), nevertheless the model's failure to match some IRFs and cross-correlations presumably accounts for its failure on the Wald statistic.

The variance decomposition is dominated by the shocks to consumer preferences, with the investment shock contributing also to Q and investment itself. Productivity and monetary shocks have little effect, underlining the model's New Keynesian provenance.

	Actual Estimate	Lower Bound	Upper Bound	State	t-stat*
A_C^C	0.88432	0.67784	2.17487	TRUE	-1.29699
A_C^I	-0.09612	-1.01321	0.71139	TRUE	0.13541
A_C^π	0.01867	-0.02040	0.20112	TRUE	-1.18651
A_C^Y	0.07935	-0.06675	1.26126	TRUE	-1.43262
A_C^R	-0.00824	-0.05480	0.19518	TRUE	-1.24489
A_I^C	-0.02461	-0.08517	0.12605	TRUE	-0.89331
A_I^I	0.91856	0.80536	1.04445	TRUE	-0.22981
A_I^π	-0.01074	-0.01390	0.01436	TRUE	-1.61228
A_I^Y	-0.01190	-0.06222	0.11812	TRUE	-0.83873
A_I^R	-0.00504	-0.01477	0.01963	TRUE	-0.85038
A_π^C	-0.04105	-2.75549	1.36838	TRUE	0.58582
A_π^I	-0.71538	0.00212	4.65028	FALSE	-2.81450
A_π^π	0.68194	0.33816	0.94001	TRUE	0.05555
A_π^Y	-0.00692	-2.00221	1.50416	TRUE	0.27295
A_π^R	-0.01605	-0.37265	0.34662	TRUE	-0.00606
A_Y^C	0.21989	-1.42637	0.58019	TRUE	1.21983
A_Y^I	0.38855	-1.10769	1.14039	TRUE	0.64902
A_Y^π	0.05457	-0.20815	0.08309	TRUE	1.54982
A_Y^Y	0.93795	-0.55498	1.18992	TRUE	1.32594
A_Y^R	0.06281	-0.22013	0.11210	TRUE	1.43788
A_R^C	-0.37666	-1.83366	0.58254	TRUE	0.40611
A_R^I	-0.97612	-3.20523	-0.88981	TRUE	1.76556
A_R^π	-0.05704	-0.22694	0.09598	TRUE	-0.04928
A_R^Y	-0.40669	-1.84650	0.09677	TRUE	0.85200
A_R^R	0.89695	0.60858	0.98785	TRUE	0.91781
	Wald Statistic	100		Joint t-test	75.7

*t-stat from bootstrap mean

Table 1: VAR Parameters & Model Bootstrap Bounds (SW model with SW rhos and variances)

	Consumption	Investment	Inflation	Output	Interest Rate
Actual	5.4711	37.1362	0.2459	3.6085	0.3636
Lower	3.2602	10.5080	0.0453	1.8286	0.1013
Upper	16.4491	101.1404	0.2578	8.6823	0.5723
Mean	7.8083	40.1094	0.1164	4.3093	0.2565

Table 2: Variance of Data and Bootstraps for SW's Model with SW rhos and variances

	Prod	Cons	Gov	Inv	Price	Int	Wage	TOTAL
C	0.561	95.276	0.293	2.577	0.092	1.173	0.027	100
I	0.622	79.123	0.036	18.336	0.161	1.689	0.034	100
K	0.499	80.629	0.054	17.609	0.098	1.084	0.028	100
L	3.472	91.455	1.435	2.028	0.105	1.454	0.051	100
π	0.295	95.017	0.033	0.310	3.285	0.880	0.181	100
Q	1.059	51.598	0.007	38.638	0.632	7.973	0.093	100
R	0.447	98.729	0.042	0.325	0.070	0.365	0.021	100
rk	0.546	90.329	0.221	7.749	0.143	0.800	0.212	100
W	0.638	86.600	0.066	7.464	0.699	2.445	2.088	100
Y	0.891	90.848	1.139	4.976	0.159	1.943	0.044	100

Table 3: Variance Decomposition for SW's Model with SW rhos and variances

More details can be found in Annex Alph@section.

We also find that the SWNC model does badly under these error assumptions — see Annex Alph@section. Thus we can summarise this test as confirming the relatively good model performance that SW themselves found with their model under the error posterior assumptions they made, while rejecting the alternative NC version we have created.

4.2 The New Classical version of SW using error properties chosen to suit NC priors

A key claim of Smets and Wouters is that the good performance of their model reflects the importance of price and wage stickiness. In this section we look at a version of SW's model with flexible wages and prices in which we add an information lag for labour supply as in the Lucas (1972) original 'islands' model. In some respects the same idea has been picked up in 'sticky' information; the difference is that the 'stickiness' of a short lag is solely due to the physical availability (via collection and publication typically) of macro data and not e.g. to 'rational inattention' or other processing costs — cf Mankiw and Reis (2002) and Sims (2003). Such a model with a one-period information lag would correspond to the original ideas of 'New Classical' macroeconomics in which prices and wages were assumed to be flexible subject to available information.

Like Smets and Wouters we assume a set of priors about the error properties but, in contrast, we assume processes in keeping with New Classical thinking. As a result of these changes our NC alternative is somewhat handicapped. Therefore, apart from our adaptation to NC form, we adopt SW's parameters and assume a simpler Taylor Rule of the form $R_t - E_t\pi_{t+1} = 1.5(\pi_t - \pi^*)$ that better stabilises the NC model than the original SW rule. However it turns out that the key assumptions are about the errors. We followed many authors of Real Business Cycle models in assuming zero shocks to consumer and investment preferences; otherwise we simply used the actual shocks generated by the data and the NC model to find both their variance and their AR parameters (details of these actual errors are given below), scaling the errors by only 0.7.

The remarkable thing about this exercise is that this SWNC version of the model is also reasonably compatible with the data. It seems that if one chooses suitable error properties one can match both NK and NC versions of the model! We find that the data variances of our key variables all lie inside the model's 95% bounds. Then the SWNC model passes the joint t-test with a value of 85.0, with only one VAR coefficient lying outside the 95% bounds. The model's variance decomposition is now dominated for real variables by productivity and labour supply (wage) shocks; and for nominal variables by labour supply and monetary (inflation) shocks. The

14

VAR IRFs for real variables under the productivity and labour supply shocks lie outside the model 95% bounds — the model underestimates them. For nominal variables they are close for these shocks and inside for the monetary shock. The cross-correlations all lie inside the 95% model bounds, with the exceptions of the cross-correlations with output of both inflation and interest rates which are somewhat more positive than the model bounds. (Nevertheless these failures are sufficient, as with SWNK above, to ensure the SWNC model's rejection by the Wald statistic.)

Thus again we find that SWNC fits reasonably well when the errors are chosen using NC priors; furthermore the SWNK fits badly under these NC priors — see Annex Alph@section.

	Actual Estimate	Lower Bound	Upper Bound	State	t-stat*
A_C^C	0.88432	0.38922	1.00516	TRUE	0.99529
A_C^I	-0.09612	-0.83115	0.13808	TRUE	1.11278
A_C^π	0.01867	-0.15051	0.22081	TRUE	-0.09127
A_C^Y	0.07935	-0.31773	0.33397	TRUE	0.42883
A_C^R	-0.00824	-0.25663	0.33100	TRUE	-0.24559
A_I^C	-0.02461	-0.06432	0.06105	TRUE	-0.60798
A_I^I	0.91856	0.84056	1.02208	TRUE	-0.43247
A_I^π	-0.01074	-0.03590	0.03334	TRUE	-0.58558
A_I^Y	-0.01190	-0.04825	0.07188	TRUE	-0.70474
A_I^R	-0.00504	-0.05913	0.05865	TRUE	-0.19061
A_π^C	-0.04105	-0.11680	1.21684	TRUE	-1.87390
A_π^I	-0.71538	0.39292	2.45209	FALSE	-4.10970
A_π^π	0.68194	0.59311	1.31572	TRUE	-1.56029
A_π^Y	-0.00692	-0.13882	1.09768	TRUE	-1.62757
A_π^R	-0.01605	-0.27846	1.01944	TRUE	-1.15563
A_Y^C	0.21989	-0.09905	0.43690	TRUE	0.37025
A_Y^I	0.38855	-0.03180	0.76454	TRUE	0.12325
A_Y^π	0.05457	-0.18409	0.13951	TRUE	0.90608
A_Y^Y	0.93795	0.54713	1.06406	TRUE	0.81169
A_Y^R	0.06281	-0.30796	0.22601	TRUE	0.67308
A_R^C	-0.37666	-1.23808	-0.19850	TRUE	1.19908
A_R^I	-0.97612	-2.70167	-0.90465	TRUE	1.84734
A_R^π	-0.05704	-0.40124	0.21020	TRUE	0.21640
A_R^Y	-0.40669	-1.12134	-0.07573	TRUE	0.80559
A_R^R	0.89695	0.02191	1.00569	TRUE	1.41461
	Wald Statistic	100.0		Joint t-test	85.0

*t-stat from bootstrap mean

Table 4: VAR Parameters & Model Bootstrap Bounds (SWNC Model with NC Priors)

	Consumption	Investment	Inflation	Output	Interest Rate
Actual	5.4711	37.1362	0.2459	3.6085	0.3636
Lower	0.6578	4.8250	0.0986	0.5769	0.1910
Upper	6.4263	56.6121	0.3985	5.3142	0.6991
Mean	2.3265	20.6644	0.2055	1.9926	0.3864

Table 5: Variance of Data and Bootstraps for SWNC Model with NC Priors

	Prod	Gov	Monetary	Labour Supply	TOTAL
C	35.621	0.388	0.032	63.959	100
I	30.511	0.240	0.001	69.248	100
K	36.036	0.141	0.000	63.823	100
L	25.702	2.225	0.046	72.026	100
π	2.543	0.643	84.751	12.064	100
Q	18.231	1.804	0.759	79.206	100
R	8.928	2.185	46.138	42.749	100
rk	32.623	0.686	0.010	66.681	100
W	83.989	0.163	0.002	15.846	100
Y	34.674	1.190	0.025	64.111	100

Table 6: Variance Decompositon for SWNCl Model with NC Priors

4.3 The actual disturbances implied by the model compared with the assumed errors

So far we have supplied both the NK and NC versions of the SW model with essentially imaginary error properties, chosen by assumption. We now turn to the actual errors derived from using the observed data. We calculate the model's structural errors, that is the error in each equation given the actual data and the expected variables in it.

Under our procedure the exact way to derive these structural residuals is to back them out of the equations and the data. For certain equations expected variables do not enter and so the structural residuals can be backed out directly from the equations and the data. However, when expected variables enter an equation, these have to be derived from the model in order to back out the residuals.

The exact way to derive these is to solve the model conditional on available information each period; but in doing so we must include in the model the time-series processes governing the errors. However, we have to estimate these time-series processes. This introduces an inexact element since in order to estimate them we need to know what the structural residuals are; but we cannot know these without knowing the time-series processes already.

One way to solve this problem is to iterate to a fixed point between the error processes used to project the model expectations and the resulting errors found when these expectations are used to calculate them. It is however possible that there is no unique fixed point, given that we do not know the order of the ARMA process governing the errors. Thus under this method we could find the wrong residuals unless we can be confident there is a unique fixed point.

An alternative procedure is to use the fact that under the null hypothesis the data will be driven by the structural model. This model inclusive of the unknown error processes has a VAR representation which we can and have estimated directly and efficiently as our auxiliary equation. This estimated VAR will when projected provide an efficient estimate of the rational expectation of

future variables. The residuals can then be calculated given these estimates. Thus instead of trying to estimate the error processes jointly with the expectations we estimate the expectations directly. Having done so we can back out the residuals and directly estimate their time-series processes. The resulting residuals will now contain an extra random error equal to the expectations error (due to the difference between estimated VAR and the true VAR implied by the structural model and the true error processes) but we can reasonably suppose that this will be small relative to the scale of the true residual itself.

This is the method we have used here, in preference to the iterative alternative above. Some experiments with this alternative produced implausibly wild residuals, presumably because of this interaction of the error processes with the expectations.

It should also be noted that we excluded the first 20 error observations from the sample because of extreme values; we also smoothed two extreme error values in Q. Thus our sample for both bootstraps and data estimation was 98 quarters, i.e. 1975q1–1999q2.

The results of our procedure are shown in Figure 2. There are six behavioural errors: consumption, investment, productivity, interest rates (monetary policy), wage- and price-setting, and one exogenous process, government spending, which only enters into the goods market-clearing equation (or 'GDP identity'). The first error is that of the Euler equation and has a standard error of 0.5%, roughly half as much again as assumed by SW (see Canzoneri et al. (2007) on the peculiarities of actual Euler equation errors), that for investment in the second has a standard error of 1.2%, around ten times that assumed by SW. Furthermore the AR coefficients (ρs) of the structural residuals are very different; there is hardly any persistence in the estimated residuals for consumption and investment, unlike the high persistence assumed by SW. In contrast, the actual inflation and Taylor Rule errors are persistent and do not have zero ρs, as assumed. The Table below shows the comparison between SW's assumed shocks and those shown in the graphs below. These differences will turn out to be an important factor in the tests we will carry out.

Variances	Cons	Inv	Inflation	Wage	Gov	Prod	Tayl rle
Data var	0.260	1.520	0.0007	0.278	0.141	0.091	0.227
SW var	0.088	0.017	0.026	0.081	0.108	0.375	0.017
Ratio	2.95	89.41	0.03	3.43	1.31	0.24	13.35
ρ							
Data	-0.101	0.063	0.154	-0.038	0.751	0.940	0.565
SW	0.886	0.917	0	0	0.956	0.828	0

Table 7: Variances of innovations and AR Coefficients (rhos) of shocks (data-generated v. SW assumed shocks)
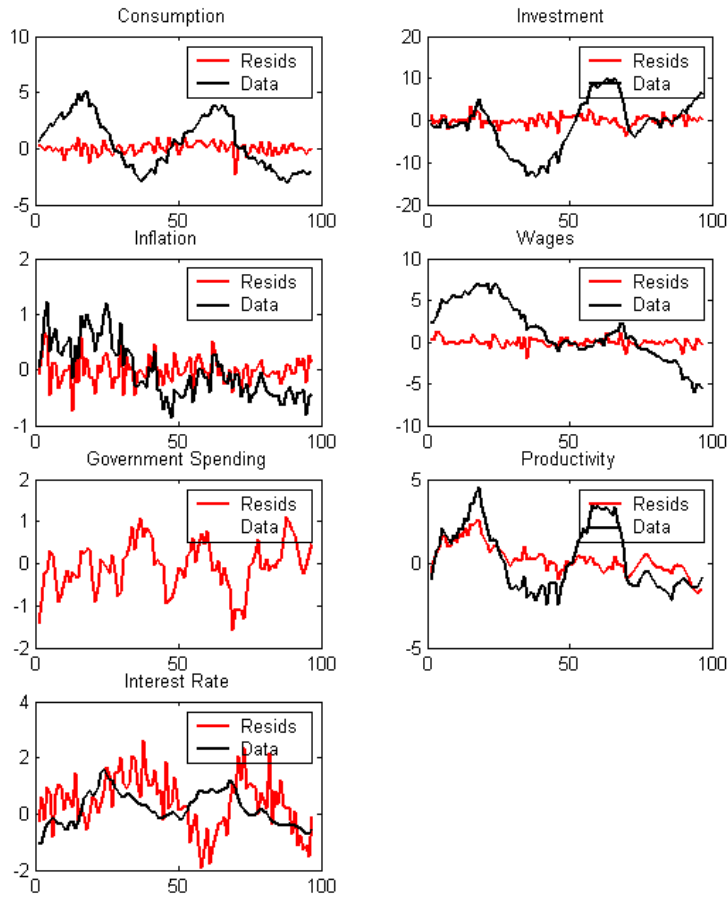
Figure 2: Single Equation Errors from SW Model

Using the actual errors we proceeded to bootstrap their model again.[1] The resulting variables'
bootstrap variances are massively in excess of the variances of the observed data — about 16
times larger. Thus the model fails to replicate the data in a rather basic way. Such findings are
not uncommon in models with calibrated parameters especially where, as here, tight bounds are
placed on the model variances to adjust them to match, say, the overall variance of GDP. Under
such a procedure it would be open to the authors to assume that the latent errors simply need to
be scaled down — this was what we did above with the assumed error properties.

The problem is more serious when the structural residuals are used. The inability of the model
to capture the scale of the variances of the data represents a real model failure. In order to take
further tests seriously we need to have a means to set the poor variance fit on one side. We do

---

[1]Needless to say, none of the above discussion on estimating the structural residuals has anything to do with
how the model is solved for given shocks once the residuals and their time-series processes have been estimated and
built into the model. The solution method here follows the standard dynare procedure.

so by notionally dividing the effect of the model disturbances into two parts: the impact effect and the effect through lags of both the errors and the endogenous variables. We could think of the model's failure as reflecting the excessive size of the impact effect. We could then scale down this impact effect, assuming that some misspecification could be isolated that is responsible for its excessive size. We can then investigate the models' lag effects or 'transmission process' — as summarised in the VAR coefficients. We note that the impact effect results from the whole model structure and whatever misspecification is responsible will also affect the lag transmission process. Nevertheless we suppose for purposes of further investigation alone that it would be possible to modify the impact effect while maintaining the transmission process.[2] Our scaling procedure is the same for all errors. We choose the scaling factor separately for each model version in such a way that the 95% bounds on the bootstraps contain the variances of the data. For the SW model as above the scaling factor needed is 0.25. In the context of SW's methods we can think of this as choosing lower error standard deviations by this amount.

4.4   The SW New Keynesian model using actual errors

Here we take the structural residuals implied by the data and base the model's behaviour on these. We then re-estimate the error processes with new autoregressive parameters (ρs). Bootstrapping their random components — drawing them as vectors to preserve any dependence between them — we get the results in Table 8 below. We have scaled the errors by 0.25; although the variances of the errors are much higher than SW's, their autocorrelation is far less so that the two effects cancel out in terms of the model's simulated variances.

The results are rather mixed. The model passes the joint t-test with a value of 93.6. However, it cannot reproduce the data variances, even with the heavy scaling we have used (Table 9). For interest rates the data variance lies very far above the model's upper bound; while for investment it lies rather below the model's lower bound. Out of the four VAR coefficients that lie outside their 95% bounds, three concern interest rate effects, as one might expect from the model's failure to capture the variance of the interest rate. The fourth concerns investment, for which the model variance is greatly excessive.

---

[2]Let the model be given by $Ay_t = ME_t y_{t+1} + N y_{t-1} + u_t$, where $u_t = \Phi u_{t-1} + \epsilon_t$. This can be transformed into $y_t = A^{-1} M B^{-1} y_t + A^{-1} N L y_t + A^{-1} u_t$; where L is the lag operator and B^{-1} is the forward operator leading the variable while keeping the date of expectations constant (here at t). Assume that the model satisfies the saddlepath Blanchard-Kahn conditions (with f forward and l backward roots), then we can rewrite it as $\prod_{i=1}^{f}(1 - \gamma_i B^{-1}) \prod_{j=1}^{l}(1 - \lambda_j L) y_t = K(L; M, N, A) u_t$. Here we note that K is a function of all the parameters of the model, as well as involving lags of the errors produced by the backward roots and current values of the errors produced by the forward roots. We can solve for y_t in terms of the current shocks and its own lagged values by projecting the forward roots onto the errors and then projecting all the backward roots, as well as error autoregressive roots, onto y_t. It is clear that the impact effect, just like the transmission effect, comes from the complete parameter set.

	Actual Estimate	Lower Bound	Upper Bound	State	t-stat*
A_C^C	0.88432	-0.72761	1.56433	TRUE	0.72681
A_C^I	-0.09612	-3.67952	1.97243	TRUE	0.52460
A_C^π	0.01867	-0.25012	0.15093	TRUE	0.63888
A_C^Y	0.07935	-1.18984	0.97379	TRUE	0.32915
A_C^R	-0.00824	-0.10728	0.07514	TRUE	0.20002
A_I^C	-0.02461	-0.12365	0.02746	TRUE	0.49674
A_I^I	0.91856	0.59652	0.96120	TRUE	1.49397
A_I^π	-0.01074	-0.01690	0.00811	TRUE	-1.04404
A_I^Y	-0.01190	-0.08305	0.05315	TRUE	0.02955
A_I^R	-0.00504	-0.00099	0.00983	FALSE	-3.34577
A_π^C	-0.04105	-0.63612	2.05088	TRUE	-1.17884
A_π^I	-0.71538	1.97340	9.92243	FALSE	-3.58523
A_π^π	0.68194	0.54220	1.11755	TRUE	-1.41208
A_π^Y	-0.00692	-0.36242	2.13874	TRUE	-1.51517
A_π^R	-0.01605	-0.08937	0.11765	TRUE	-0.53770
A_Y^C	0.21989	-0.81029	1.85980	TRUE	-0.36596
A_Y^I	0.38855	-2.35381	4.09609	TRUE	-0.23243
A_Y^π	0.05457	-0.16908	0.27962	TRUE	0.03882
A_Y^Y	0.93795	-0.32156	2.14998	TRUE	0.09489
A_Y^R	0.06281	-0.11977	0.08366	TRUE	1.50196
A_R^C	-0.37666	-4.28405	-0.20281	TRUE	1.75415
A_R^I	-0.97612	-7.78487	1.53307	TRUE	0.89647
A_R^π	-0.05704	-0.65743	-0.00702	TRUE	1.62513
A_R^Y	-0.40669	-4.19901	-0.44180	FALSE	1.99106
A_R^R	0.89695	0.45640	0.75317	FALSE	3.78248
	Wald Statistic	100.0		Joint t-test	93.6

*t-stat from bootstrap mean

Table 8: VAR Parameters & Model Bootstrap Bounds (SW with estimated rhos)

	Consumption	Investment	Inflation	Output	Interest Rate
Actual	5.4711	37.1362	0.2459	3.6085	0.3636
Lower	3.5651	47.5781	0.0745	3.3518	0.0244
Upper	23.6006	409.0063	0.3819	22.1150	0.0808
Mean	10.2900	168.5409	0.1857	9.8493	0.0442

Table 9: Variance of Data and Bootstraps for SW's Model (with estimated rhos)

	Prod	Cons	Gov	Inv	Price	Int	Wage	TOTAL
C	2.875	0.163	0.007	0.056	0.033	96.845	0.021	100
I	1.361	0.000	0.003	0.586	0.037	97.995	0.017	100
K	2.521	0.000	0.004	0.376	0.034	97.044	0.021	100
L	2.612	0.161	0.296	0.063	0.031	96.806	0.032	100
π	0.651	0.002	0.003	0.010	1.739	97.426	0.168	100
Q	0.140	0.005	0.002	0.075	0.042	99.724	0.012	100
R	6.430	1.099	0.190	0.137	0.399	91.511	0.234	100
rk	1.333	0.047	0.077	0.255	0.074	97.977	0.236	100
W	2.515	0.003	0.002	0.060	0.114	96.580	0.727	100
Y	2.565	0.108	0.196	0.077	0.034	97.001	0.020	100

Table 10: Variance Decompositon for SW model (with estimated rhos)

Turning to the VAR IRFs it can be seen (Annex Alph@section) that rejections are scattered across the five variables in our VAR and also across the structural shocks. In particular, we find that the interest rate responses are well outside their permitted bounds for most shocks. This is consistent with the model's failure to capture the variance of the interest rate. We show a number of these below.
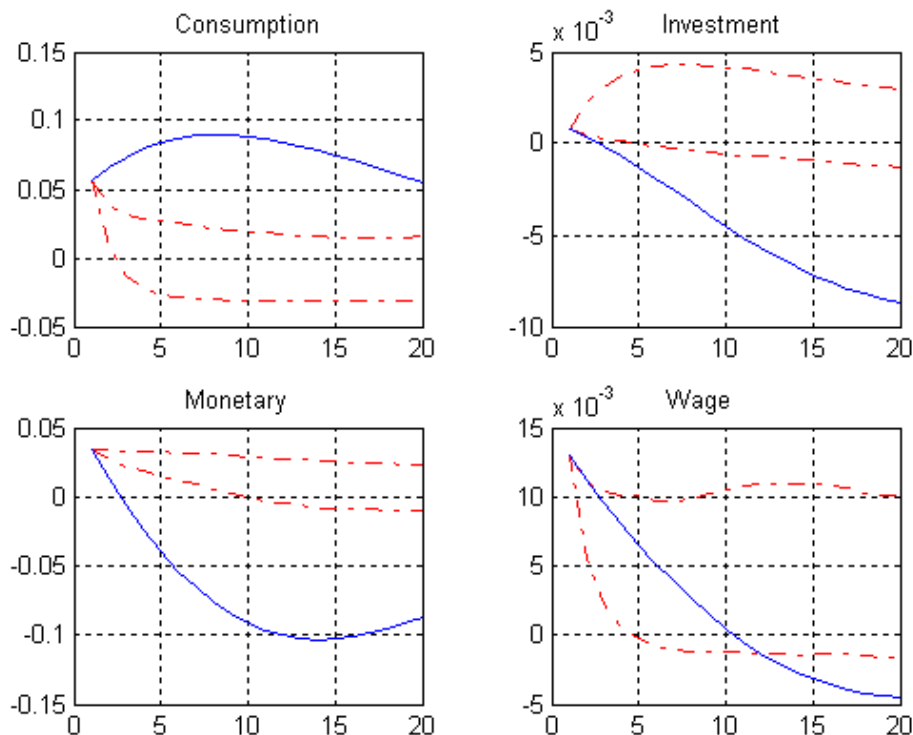

Figure 3: Interest Rate Responses to Various Shocks with 95% Bounds (SWNK with actual errors)

The cross-correlations (see Annex Alph@section) reveal that the model underpredicts the autocorrelation of consumption and badly underpredicts the positive data cross-correlations of both consumption and investment with output (both from lagged output to consumption and investment and from lagged consumption and investment to output).

It is therefore not surprising that this model is rejected on the Wald statistic. What seems to be undermining the model's performance is the virtual elimination of the autocorrelation in both the main demand shocks, consumption and investment. This is what generates the poor prediction of the persistence and cross-correlations for the real variables. It also produces too little interest rate variation because inflation and so interest rates respond less to less persistent shocks. The overwhelmingly dominant shock is now that to the Taylor Rule, because it has a high variance (it includes all the shocks to potential output) as well as moderate persistence.

21

Thus in summary the SWNK model fails in several important ways to match the data, once actual error behaviour is substituted for assumed behaviour. The reason is the drop in persistence of the actual demand shocks compared with SW's assumed shocks.

4.5   The New Classical Version of SW using actual errors

Here we used the same SWNC model as above but with the full set of actual errors. The model now requires scaling by 0.67. It passes the joint t-test with a value of 76.8. However like SWNK this too cannot match the individual variable variances well. It overpredicts both inflation and interest rate variances both of which lie rather below the model's lower 95% bound. This is reflected in the four VAR coefficients that lie outside their bounds: the cross-effects from inflation and interest rates to consumption and from inflation on interest rates, all of which the model makes excessively negative, and the partial autocorrelation of interest rates which the model greatly underpredicts. The excessive variation in inflation and associated interest rates produces low interest rate persistence; and the model generates from it high negatively correlated responses of consumption.

The same story shows up in the IRFs. The dominant shocks on real variables are now the labour supply (wage) and productivity shocks while for nominal variables they are consumption, monetary and labour supply. We find that the responses of consumption and output to labour supply and productivity lie well outside the rather narrow model 95% bounds, reflecting the model's inadequate variance for these. Those for interest rates and inflation start in or close to the bounds; but thereafter they die off more slowly in the data than in the model — see Figure 4

	Actual Estimate	Lower Bound	Upper Bound	State	t-stat*
A_C^C	0.88432	0.52259	1.18525	TRUE	0.12292
A_C^I	-0.09612	-0.96369	1.27543	TRUE	-0.42842
A_C^π	0.01867	-0.97493	-0.01531	FALSE	1.96524
A_C^Y	0.07935	-0.17288	0.45474	TRUE	-0.36065
A_C^R	-0.00824	-1.23870	-0.03653	FALSE	1.97025
A_I^C	-0.02461	-0.06602	0.04563	TRUE	-0.69450
A_I^I	0.91856	0.68911	1.07953	TRUE	0.19148
A_I^π	-0.01074	-0.07240	0.09956	TRUE	-0.49211
A_I^Y	-0.01190	-0.04235	0.05387	TRUE	-0.76602
A_I^R	-0.00504	-0.08904	0.12880	TRUE	-0.37198
A_π^C	-0.04105	-0.94409	1.02284	TRUE	-0.29536
A_π^I	-0.71538	-2.79046	4.28830	TRUE	-0.92632
A_π^π	0.68194	0.16140	3.20631	TRUE	-1.45758
A_π^Y	-0.00692	-0.84830	1.03652	TRUE	-0.30382
A_π^R	-0.01605	-0.60757	3.17618	TRUE	-1.52104
A_Y^C	0.21989	-0.31978	0.50850	TRUE	0.56298
A_Y^I	0.38855	-1.38295	1.34478	TRUE	0.53247
A_Y^π	0.05457	-0.35789	0.84976	TRUE	-0.60112
A_Y^Y	0.93795	0.38835	1.13783	TRUE	0.85587
A_Y^R	0.06281	-0.39085	1.05363	TRUE	-0.64525
A_R^C	-0.37666	-0.81912	0.73652	TRUE	-0.80463
A_R^I	-0.97612	-3.31798	2.22810	TRUE	-0.18695
A_R^π	-0.05704	-2.57757	-0.25199	FALSE	2.37433
A_R^Y	-0.40669	-0.78257	0.66123	TRUE	-0.81601
A_R^R	0.89695	-2.64767	0.23123	FALSE	2.91387
	Wald Statistic	100.0		Joint t-test	76.8

*t-stat from bootstrap mean

Table 11: VAR Parameters & Model Bootstrap Bounds (SWNC Model)

	Consumption	Investment	Inflation	Output	Interest Rate
Actual	5.4711	37.1362	0.2459	3.6085	0.3636
Lower	0.7330	9.8999	0.4497	0.6434	0.6906
Upper	6.8787	84.0695	0.9583	5.6327	1.4368
Mean	2.4783	34.5926	0.6678	2.2226	1.0270

Table 12: Variance of Data and Bootstraps for SWNC Model

	Prod	Cons	Gov	Inv	Price	Wage	TOTAL
C	35.360	0.237	0.385	0.495	0.032	63.491	100
I	27.903	0.060	0.219	8.488	0.001	63.329	100
K	34.697	0.027	0.135	3.689	0.000	61.451	100
L	25.517	0.278	2.209	0.442	0.046	71.508	100
π	1.739	31.512	0.440	0.083	57.973	8.252	100
Q	15.320	13.421	1.516	2.547	0.638	66.558	100
R	5.607	36.900	1.372	0.301	28.975	26.846	100
rk	31.196	0.090	0.656	4.285	0.010	63.763	100
W	83.085	0.022	0.161	1.053	0.002	15.675	100
Y	34.437	0.152	1.182	0.531	0.025	63.673	100

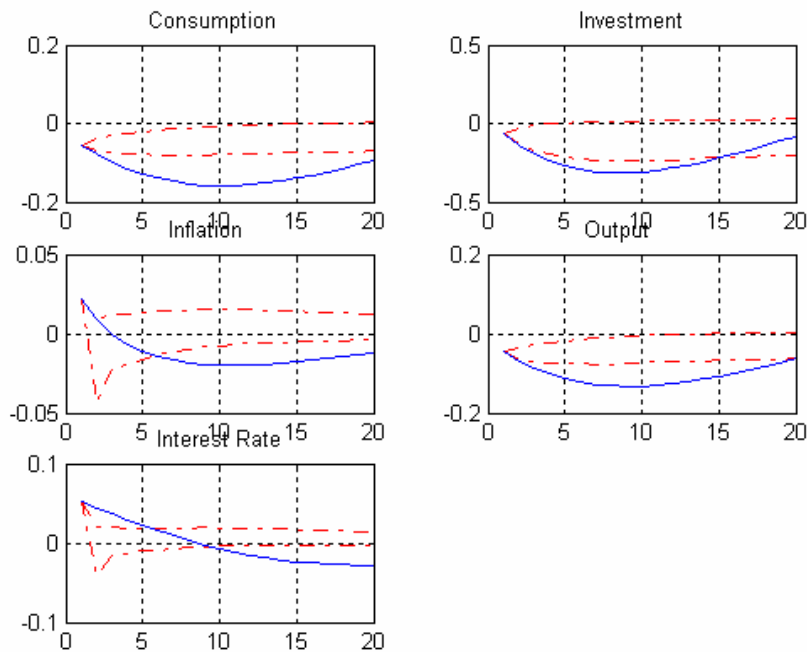Table 13: Variance Decompositon for SWNC Model

Figure 4: Labour Supply Shock to New Classical Version of SW

We can see the same story too in the cross-correlations. Here the autocorrelations of consumption and output, and also their cross-correlations are picked up well by the model; even though it underpredicts their variance, it has no trouble with their dynamic patterns. However, for inflation and interest rates the model greatly underpredicts their autocorrelations, and also their cross-correlations with output. The excessive variance of both the former due to the model's price-flexibility does not produce enough correlation with persistent output.

So in summary the new classical version of the model too fails to match the data — and is also rejected on the Wald statistic. Here the reason is that the substantial demand shocks produce high variance in inflation and interest rates; because these shocks die out quickly they also produce too little persistence to match the data for these.

If we compare the New Classical with the New Keynesian versions of SW we may note that in SWNC supply shocks determine real variables while demand shocks determine nominal ones; whereas in SWNK demand shocks dominate both real and nominal variables. SW's original NK model assumed highly persistent demand shocks and these were helpful in matching the data's dynamic patterns. However, withdraw the persistence of these shocks, as we find in the actual errors, and the NK model fails — essentially because it cannot generate enough persistence in real variables. As for the NC version, these demand shocks retain high variance together with low persistence; and this produces the same patterns in nominal variables which now fail to match the

data. Real variables are well-matched by the model's response to persistent supply shocks.

4.6 Robustness tests treating the data in alternative ways

We have placed much emphasis on the actual errors implied by the model and the data. This raises the issue of whether the errors are affected in any important way by alternative ways of treating the data. One major issue is that of filtering the data to obtain approximate stationarity. SW simply take out a constant and linear trend. However, this leaves one data-series on the borderline of trend-stationarity: viz wages, though this is not included in our VAR while its error has virtually no autocorrelation. We redid the exercise using the Hodrick-Prescott filter.

Another question that arises is whether the data were affected by non-stochastic elements which acted in the same way as 'trend' factors. Thus we could attempt some division of the errors into 'non-stochastic' (once-for-all events) and stochastic (repetitive events). This is similar to the idea of detrending in the sense that it removes elements that are not properly components of the business cycle. To do this accurately would require us to specify what these once-for-all events were and remove their effects both from the model and from the data, just as in principle we do for 'trend'. Such one-off events over the EU's history include waves of new membership and German reunification. In another check we redid the exercise after regressing the SW data on a set of dummies for idnetifiable events of this sort — we then took these dummy effects out of the data and used the adjusted data from that point.

Results of both these checks are reported in our annexes from Annex Alph@section to Annex Alph@section; they make no material difference to our results.

5 Comparing the New Keynesian and New Classical versions of the SW model — a weighted combination?

In this paper we have sought to test the SW model and to assess its capacity to replicate dynamic features of the data as compared with a quite different model. Such a comparison makes sense if there is still significant controversy about what type of model should be used, and it matters both for understanding events and for making policy. SW made certain prior modelling choices, some of which remain controversial within macroeconomics. One of these was their main innovation, the assumptions relating to the degree of price and wage rigidity in their model. We have re-examined the implications of these assumptions at some length by positing a New Classical version of their DSGE model which does not have price and wage stickiness.

We found that the properties of the errors are the key element in the success or failure of both SWNK and SWNC in these tests. The more the error properties conform to NK priors, with dominant demand shocks, the better the SWNK model performs and the worse the SWNC does. In contrast, the more the errors conform to New Classical priors, the better the SWNC performs and the worse SWNK does. When the error properties are derived from observed data, both models have difficulty fitting the data, though SWNC model is probably the closest to doing so. What is the explanation for these results?

In the SWNK model, because capacity utilisation is flexible, demand shocks (consumption/ investment/money) dominate output and — via the Phillips Curve — inflation, then — via the Taylor Rule — interest rates. Supply shocks (productivity, labour supply, wages/inflation mark-ups) play a minor role as 'cost-push' inflation shocks as they do not directly affect output. Persistent demand shocks raise 'Q' persistently and produce an 'investment boom' which, via demand effects, reinforces itself. Thus the model acts as a 'multiplier/accelerator' of demand shocks. Demand shocks therefore dominate the model, both for real and nominal variables. Moreover, in order to obtain good model performance for real and nominal data, these demand shocks need to be of sufficient size and persistence.

In the SWNC model an inelastic labour supply causes output variation to be dominated by supply shocks (productivity and labour supply) and investment/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 SWNC 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 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.

The observed demand shocks have too little persistence to capture the variability of real variables in the SWNK model, but they generate too much variability in nominal variables in the SWNC model. The observed supply shocks matter little for the SWNK but are about right in size and persistence for the real variables in the SWNC. The implication is that the flexibility of prices and wages may lie somewhere between New Keynesian and the New Classical models. For example, adding a degree of price and wage stickiness to the SWNC model would bring down the variance of nominal variables, and boost that of real variables in the model.

A natural way to look at this is to assume 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 can assume that the size of each sector depends on the facts of competition and do not vary in our sample. The degree of imperfect competition could differ between labour and product markets. For the exercise here we will initially assume that it is the same in each market and given by a single free parameter, v. This implies that the price and wage equations will be a weighted average of the SWNK and SWNC equations, with the weights respectively of $(1 - v)$ and v. We will also assume that the monetary authority uses this parameter to weight its New Keynesian and New Classical Taylor Rules as we have found that different rules work best for a competitive (NC) model and an imperfect competition (NK) economy. In practice we can think of the weight v as giving the extent of the NC (competitive) share of the economy.

We now choose a value of v for which the combined model is closest to matching the data variances while also passing the joint t-test. This is an informal use of indirect inference which provides a broader criterion which better reflects our concerns with the models' perfomance than simply applying a Wald score to, for example, the VAR coefficients. The optimal value turns out to be 0.94. This implies quite a small NK sector of only 6% of the economy, but it is sufficient to bring the overall economy's properties close to the dynamic facts. We allowed the weight to be further varied around this to generate an optimum performance: in labour markets ($v_w = 0.08$), product markets ($v_p = 0.06$), and monetary policy ($v_m = 0.04$). We now consider how good a fit this is.

The key difference is the ability of the model to replicate the variances in the data. No scaling of the shocks is required and all the data variances lie within the model's 95% bounds (Table 15). The model therefore satisfies the necessary basic conditions for us to take it seriously: it produces behaviour of the right size for both real and nominal variables when the structural errors are generated from the model using the observed data and are not rescaled.

The joint t-test is 90.8 with just three VAR coefficients lying outside their 95% bounds. The main discrepancy is the partial autocorrelation of interest rates which the model underpredicts. The other two coefficients involve the cross-effects of inflation and interest rates on consumption, which are marginally outside their bounds.

The variance decomposition of real variables is now heavily skewed towards being caused by supply shocks with 75% of the output variance being due in the model to labour supply and productivity shocks. In contrast, nominal variables are dominated by demand shocks with 74% of the variance of inflation due in the model to the shocks to the Taylor Rule. Being the sum of a real variable and expected inflation, about one thirds of the variance of nominal interest rates is due to productivity and labour supply shocks, with virtually all the remainder due to shocks to

	Actual Estimate	Lower Bound	Upper Bound	State	t-stat*
A_C^C	0.88432	0.26572	1.02827	TRUE	1.17662
A_C^I	-0.09612	-0.92015	1.19403	TRUE	-0.39647
A_C^π	0.01867	-0.30244	0.08489	TRUE	1.27573
A_C^Y	0.07935	-0.36553	0.28805	TRUE	0.72452
A_C^R	-0.00824	-0.33625	-0.00284	TRUE	1.83460
A_I^C	-0.02461	-0.06640	0.06491	TRUE	-0.75793
A_I^I	0.91856	0.68054	1.05976	TRUE	0.30909
A_I^π	-0.01074	-0.02810	0.05323	TRUE	-1.14494
A_I^Y	-0.01190	-0.04254	0.07168	TRUE	-0.91032
A_I^R	-0.00504	-0.01937	0.04928	TRUE	-1.13906
A_π^C	-0.04105	-0.28473	1.84441	TRUE	-1.51365
A_π^I	-0.71538	-1.68976	4.13846	TRUE	-1.38882
A_π^π	0.68194	0.44617	1.57336	TRUE	-1.23414
A_π^Y	-0.00692	0.08981	1.93824	FALSE	-2.00415
A_π^R	-0.01605	0.10374	1.04049	FALSE	-2.51828
A_Y^C	0.21989	-0.27066	0.71764	TRUE	0.03638
A_Y^I	0.38855	-1.34895	1.33299	TRUE	0.50045
A_Y^π	0.05457	-0.25489	0.27741	TRUE	0.25731
A_Y^Y	0.93795	0.38837	1.25734	TRUE	0.42118
A_Y^R	0.06281	-0.19508	0.25990	TRUE	0.19567
A_R^C	-0.37666	-2.04721	0.26606	TRUE	0.87186
A_R^I	-0.97612	-4.17678	2.09513	TRUE	0.07900
A_R^π	-0.05704	-1.05222	0.07775	TRUE	1.45411
A_R^Y	-0.40669	-2.12602	-0.15480	TRUE	1.28935
A_R^R	0.89695	-0.52286	0.45459	FALSE	3.79338
Wald Statistic	100.0			Joint t-test	90.8

*t-stat from bootstrap mean

Table 14: VAR Parameters & Model Bootstrap Bounds (Weighted Model)

	Consumption	Investment	Inflation	Output	Interest Rate
Actual	5.4711	37.1362	0.2459	3.6085	0.3636
Lower	1.7200	20.7905	0.2292	1.5284	0.2036
Upper	13.7364	172.3241	0.8405	11.3359	0.7146
Mean	5.0452	69.2529	0.4425	4.4535	0.3764

Table 15: Variance of Data and Bootstraps for the Weighted Model

	Prod	Cons	Gov	Inv	Price	Lab. Supply	Wage SW	Taylor Rule	TOTAL
C	26.041	6.513	0.413	2.903	0.068	49.568	0.000	14.494	100
I	13.792	0.054	0.158	34.390	0.003	33.683	0.000	17.920	100
K	20.767	0.033	0.121	18.648	0.002	39.724	0.000	20.706	100
L	22.861	4.391	6.197	2.762	0.033	54.190	0.000	9.566	100
π	1.087	11.414	0.556	0.469	6.091	6.246	0.002	74.135	100
Q	9.344	5.299	1.284	12.582	1.428	43.131	0.001	26.931	100
R	5.141	20.320	2.297	2.421	6.155	27.080	0.011	36.574	100
rk	8.240	32.100	1.832	9.411	0.960	17.745	0.004	29.708	100
W	14.806	36.872	0.182	2.092	1.301	9.770	0.008	34.969	100
Y	24.531	4.518	3.780	3.538	0.052	48.228	0.000	15.352	100

Table 16: Variance Decompositon for Weighted Model

the Taylor Rule (a quarter) and to other demand shocks.

The model bounds for the VAR impulse response functions enclose many of the data-based IRFs for the three key shocks: labour supply, productivity and shocks to the Taylor Rule. The main discrepancies in response to supply shocks are the longer-term predictions of interest rates and inflation which in the data wander further from equilibrium than in the model. Again, apart from the longer-term interest rate predictions, all data-based IRFs for the Taylor Rule demand shock lie inside; these are (as we saw from the VAR coefficient) a lot more persistent in the data than in the model. Hence the model performance based on the IRFs is fairly good, with the main weakness in the interest rate prediction. However, the failures on the IRFs that we observe can plausibly account for the model's failure on the Wald statistic, in common with all versions of the model we have examined.
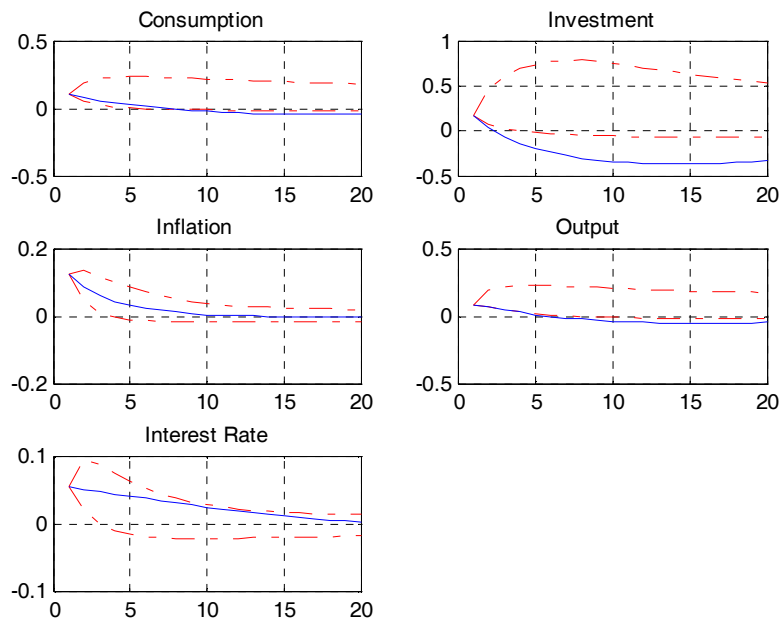

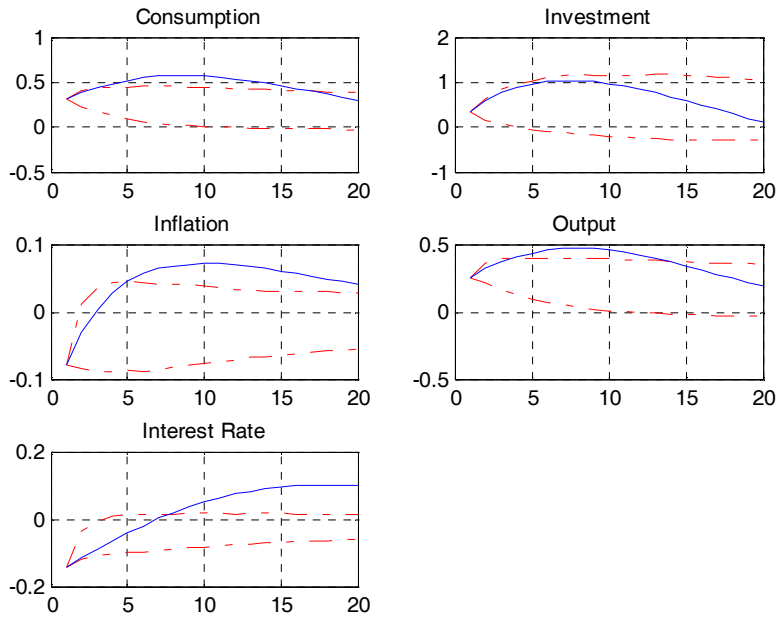Figure 5: Taylor Rule Shock to Weighted Model

Figure 6: Productivity Shock to Weighted Model

Looking at cross-correlations for the real variables, we find, as for the New Classical model alone, that the data-based correlations all lie inside the model's bounds. Now, however, the weights on SWNK wage, price and Taylor rule equations, although small, produce behaviour in the nominal variables that is almost within the 95% bounds of the weighted model (only the interest rate cross-correlation with output lies much outside). Again, this marginal failure is consistent with the Wald statistic overall rejection of the model.
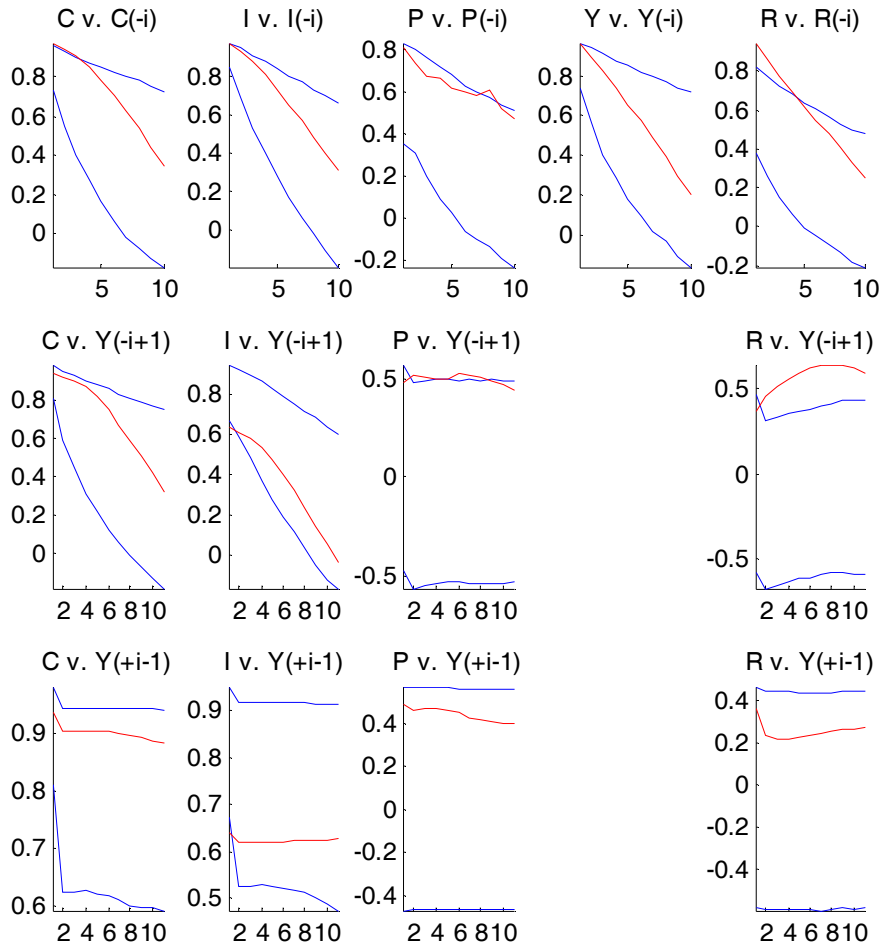
Figure 7: Cross-Correlations for weighted model (C=consumption, I=investment, P=inflation, Y=output, R=nominal interest rate)

To summarise, we find that small weights on SWNK wage, price and interest rate equations suffice to get the mixed New Classical-New Keynesian model to pass most of our tests. There are still some failures, so that the problem of finding a fully satisfactory specification remains, as indicated by the continued failure to pass the overall Wald statistic. Nonetheless, within the specifications at our disposal here, we can say that the EU economy appears to be closest to a New Classical specification.

We note that these methods could be applied to other features of Smets and Wouters's model. The method of indirect inference permits a variety of explorations of alternative modelling choices while maintaining the overarching DSGE framework. It is possible that it is simply too hard for a DSGE model to pass the tests we propose here (the viewpoint of Canova (1994) and also an early

viewpoint of Lucas and Prescott cited in Evans and Honkapohja (2005) for example). But in our experiments here the DSGE models we look at come close enough in some cases to suggest that this is too pessimistic a view.

6   Conclusion

In this paper we have applied the method of indirect inference to testing an influential DSGE model for the EU created by Smets and Wouters (2003). In many key respects this model follows an approach developed by Christiano et al. (2005) for the US. Using indirect inference a structural model's parameters may be chosen to optimise its capacity to replicate the parameters of an auxiliary time-series model whose role is to describe the data parsimoniously. Instead of using indirect inference to estimate the model, here we use it to test the model by deriving the small sample distribution of our test statistic under the null hypothesis that a calibrated or estimated structural DSGE model is correct. The DSGE model's errors are recovered and used for bootstrapping (after whitening); the resulting pseudo-samples are used to obtain the sampling distribution it implies for the parameters of the auxiliary time series model. The test then consists of determining whether functions of the parameters of the time-series (VAR) model estimated on the actual data lie within some confidence interval of this distribution. We use a joint t-test and a Wald test statistic to evaluate the overall fit of the DSGE model to the whole set of VAR parameters.

Our interest in conducting such a test lies in our wish to discriminate between very different models' capacities to embrace the dynamic behaviour of the data; in particular New Keynesian models with substantial nominal rigidity as compared with models with flexible prices. We found that the SW model in any form exhibits greatly excessive volatility compared with the data. If this misspecification is dealt with by rescaling to match the data variances, then much depends on the error properties used in the testing process. If one uses SW's assumed error properties (but rescaled) then their model fits the data reasonably well, which is consistent with their own findings, but the New Classical version fails. If, however, one uses New Classical priors, which emphasise price flexibility and supply shocks, then the New Classical version fits reasonably and SW's New Keynesian model fails. If the actual error properties are assumed then both models have difficulties fitting the data variances. The New Keynesian produces too little variation in nominal variables (especially interest rates) and too much in real variables; in contrast, the New Classical model has too much variation in nominal and too little in real variables. When the two models are artificially combined, and the weight on the NC formulation is dominant at over 90%, the combined model passes most of our tests. This suggests that only some minor modification of the NC to allow for a

degree of nominal rigidity is required in order to get close to fitting the dynamic facts. Nevertheless more work has to be done to get even this model to fit entirely satisfactorily, as ultimately tested by the overall Wald statistic.

These preliminary explorations in testing a large DSGE model raise two main issues that require further work. First, Smets and Wouters estimated their model using Bayesian methods. This relies on priors about the error distributions, whereas our classical testing procedure is based on solving the model using the observed data. This leads to different conclusions. It would be helpful to gain a better understanding of how the two approaches may be made compatible. Second, we have found faults with the SW model's structure (in all the versions examined) which suggests the need for further work on model specification.

References

Canova, F. (1994), 'Statistical inference in calibrated models', *Journal of Applied Econometrics* **9**, S123–S144.

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

Canzoneri, M. B., Cumby, R. E. and Diba, B. T. (2007), 'Euler equations and money market interest rates: A challenge for monetary policy models', *Journal of Monetary Economics* **54**, 1863–1881.

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.

Evans, G. W. and Honkapohja, S. (2005), 'An interview with thomas j. sargent', *Macroeconomic Dynamics* **9**, 561–583.

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.

Juillard, M. (2001), 'Dynare: A program for the simulation of rational expectations models', *Computing in Economics and Finance* **213**.

Lucas, R. (1972), 'Expectations and the neutrality of money', *Journal of Economic Theory* **4**, 103–24.

Mankiw, G. and Reis, R. (2002), 'Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve', *Quarterly Journal of Economics* **117**, 1295–1328.

Sims, C. (2003), 'Implications of rational inattention', *Journal of Monetary Economics* **50**, 665–690.

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

Appendix A The Smets-Wouters model

Households

The household utility function is

$$U\left(c_{it}, n_{it}, \frac{M_{it}}{P_t}\right) = \left[\frac{(c_{it} - hc_{it-1})^{1-\sigma_c}}{1-\sigma_c} - \frac{n_{it}^{1+\sigma_n}\varepsilon_t^n}{1+\sigma_n} + \frac{\left(\frac{M_{it}}{P_t}\right)^{1-\sigma_m}\varepsilon_t^m}{1-\sigma_m}\right]\varepsilon_t^B$$

where c_{it}, n_{it}, and $\frac{M_{it}}{P_{it}}$ denote the consumption, work and real money balances of the ith household, the ε_t^i, $(i = B, n, m)$ are preference shocks, and P_t is the general price level. The term hc_{t-1} is to capture consumption habits, where c_t is aggregate consumption. The real household budget constraint is

$$\frac{M_{it}}{P_t} + p_t^B \frac{B_{it}}{P_t} = \frac{M_{t-1}}{P_{t-1}} + \frac{B_{i,t-1}}{P_{t-1}} + y_{it} - c_{it} - i_{it}$$

where bonds B_{it} are one-period securities with a price of p_t^B. Total household income is

$$y_{it} = w_{it}n_{it} + a_{it} + r_t^k z_{it} k_{i,t-1} - \Psi(z_{it})k_{i,t-1} + d_{it}$$

where w_{it} is the real wage rate, k_{it} is the capital stock, r_t^k is the rate of return to capital, the term $r_t^k z_{it} k_{i,t-1} - \Psi(z_{it})k_{i,t-1}$ represents income from capital after depreciation, z_{it} is capacity utilisation and d_{it} is dividend income.

The resulting Euler equation is

$$E_t\left[\beta \frac{\lambda_{t+1}}{\lambda_t} \frac{R_t P_t}{P_{t+1}}\right] = 1$$

where R_t is the gross nominal rate of return on bonds $(R_t = \frac{1}{p_t^B})$ and λ_t is the marginal utility of consumption:

$$\lambda_t = (c_t - hc_{t-1})^{-\sigma_c}\varepsilon_t^B$$

The demand for money is

$$\left(\frac{M_t}{P_t}\right)^{-\sigma_m}\varepsilon_t^m = (c_t - hc_{t-1})^{-\sigma_c} - \frac{1}{R_t}$$

Households are assumed to act as price setters in the labour market. Their nominal wages are given by

$$W_{it} = \left(\frac{P_{t-1}}{P_{t-2}}\right)^\gamma W_{i,t-1}$$

Households set their nominal wages to maximise their inter-temporal objective function subject to their budget constraint and the demand for labour which is given by

$$n_{it} = \left(\frac{W_{it}}{W_t} \right)^{-\frac{1+\lambda_{w,t}}{\lambda_{w,t}}} n_t$$

where n_t the aggregate labour demand and W_t, the aggregate nominal wage, are given by

$$
\begin{aligned}
n_t &= \left[\int_0^1 (n_{it})^{\frac{1}{1+\lambda_{w,t}}} \, di \right]^{1+\lambda_{w,t}} \\
W_t &= \left[\int_0^1 (W_{it})^{-\frac{1}{\lambda_{w,t}}} \, di \right]^{-\lambda_{w,t}}
\end{aligned}
$$

and

$$\lambda_{w,t} = \lambda_w + \eta_t^w$$

and η_t^w is an $i.i.d.$ shock.

The result of this maximisation is the following mark-up equation for the re-optimised wage:

$$\frac{\tilde{w}_t}{P_t} E_t \Sigma_{s=0}^{\infty} \beta^s \xi_w^s \left(\frac{P_t/P_{t-1}}{P_{t+s}/P_{t+s-1}} \right)^{\gamma} \frac{n_{i,t+s} U_{c,t+s}}{1 + \lambda_{w,t+s}} = E_t \Sigma_{s=0}^{\infty} \beta^s \xi_w^s n_{i,t+s} U_{n,t+s}$$

where $\tilde{w}_t$ is the new optimal nominal wage, $\xi_w = 0$ if wages are perfectly flexible. The real wage is a mark-up $1 + \lambda_{w,t}$ over the current ratio of the marginal disutility of labour to the marginal utility of an additional unit of consumption. As a result, the aggregate wage satisfies

$$W_t^{-\frac{1}{\lambda_w}} = \xi \left[W_{t-1} (\frac{P_{t-1}}{P_{t-2}})^{\gamma} \right]^{-\frac{1}{\lambda_w}} + (1-\xi)\tilde{w}_t^{-\frac{1}{\lambda_w}}$$

Households, who own firms, choose the capital stock and investment to maximise their inter-temporal subject to their budget constraint and the capital accumulation condition

$$k_t = (1-\delta)k_{t-1} + I\left(\frac{i_t \varepsilon_t^i}{i_{t-1}} \right) i_t$$

where $I(\frac{i_t \varepsilon_t^i}{i_{t-1}})$ is an adjustment cost function, and ε_t^i is an investment shock determined by the autoregression

$$\varepsilon_t^i = \rho \varepsilon_{t-1}^i + \eta_t^i$$

The first-order conditions are

$$Q_t = E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\left[Q_{t+1}(1-\delta)+z_{t+1}r^k_{t+1}-\Psi(z_{t+1})\right]\right]$$

$$1 = Q_tI'\left(\frac{i_t\varepsilon^i_t}{i_{t-1}}\right)\left(\frac{i_t\varepsilon^i_t}{i_{t-1}}\right)+\beta E_tQ_{t+1}\frac{\lambda_{t+1}}{\lambda_t}\left(\frac{i_{t+1}\varepsilon^i_{t+1}}{i_t}\right)\left(\frac{i_{t+1}\varepsilon^i_{t+1}}{i_t}\right)\left(\frac{i_{t+1}}{i_t}\right)$$

$$r^k_{t+1} = \Psi'(z_t)$$

where Q_t is the value of installed capital.

Firms

It is assumed that there is a single final competitive good and a continuum of monopolistically produced intermediate goods indexed by j, where j is distributed over the unit interval ($j \in [0,1]$). The final good is produced by

$$y_t = \left[\int_0^1 y_t^{j\frac{1}{1+\lambda_{p,t}}}dj\right]^{1+\lambda_{p,t}}$$

where y_t^j is the intermediate good and ν_t is a mark-up generated by

$$\lambda_{p,t} = \lambda_p + \eta^p_t$$

where η^p_t is an *i.i.d.* shock. Cost minimisation gives the demand function for intermediate goods as

$$y_{jt} = \left(\frac{p_{jt}}{P_t}\right)^{-\frac{1+\lambda_{p,t}}{\lambda_{p,t}}}y_t$$

and the final goods price level which is

$$P_t = \left[\int_0^1 (p_{jt})^{-\frac{1}{\lambda_{p,t}}}dj\right]^{-\lambda_{p,t}}$$

where p_{jt} are the prices of intermediate goods.

Intermediate goods are produced using the technology

$$y_{jt} = (z_tk_{j,t-1})^\alpha N_{j,t}^{1-\alpha}\varepsilon^a_t - \Phi$$

where $N_{j,t}$ is an index of different types of labour used by firms, Φ is a fixed cost and ε^a_t is the productivity shock. Cost minimisation implies that

$$\frac{W_tN_{j,t}}{r^k_tz_tk_{j,t-1}} = \frac{1-\alpha}{\alpha}$$

The firm's marginal cost is

$$MC_t = \frac{1}{\varepsilon_t^a} W_t^{1-\alpha} \left(r_t^k\right)^a \left[\alpha^{-\alpha}(1-\alpha)^{-(1-\alpha)}\right]$$

which is independent of the intermediate good produced. Nominal firm profits are

$$\pi_{j,t} = (p_{jt} - MC_t) \left(\frac{p_{jt}}{P_t}\right)^{-\frac{1+\lambda_{p,t}}{\lambda_{p,t}}} y_t - MC_t \Phi$$

Firms are assumed to be able to re-optimise their price randomly with probability $1 - \xi_p$ as in the Calvo model. The optimal price $\tilde{p}_t$ is obtained from the first-order condition

$$E_t \Sigma_{s=0}^{\infty} \beta^s \xi_p^s \lambda_{t+s} y_{j,t+s} \left[\frac{\tilde{p}_t}{P_t} \left(\frac{P_{t+s-1}/P_{t-1}}{P_{t+s}/P_t}\right)^{\gamma} - (1+\lambda_{p,t+s})\frac{MC_{t+s}}{P_{t+s}}\right] = 0$$

which shows that the optimal price is a function of future marginal costs, and is a mark-up over them unless $\lambda_p = 0$. The general price index therefore satisfies

$$P_t^{-\frac{1}{\lambda_{p,t}}} = \xi_p \left(P_{t-1}(\frac{P_{t-1}}{P_{t-2}})^{\lambda_p}\right)^{-\frac{1}{\lambda_{p,t}}} + (1-\xi_p)\tilde{p}_t^{-\frac{1}{\lambda_{p,t}}}$$

Market equilibrium

Final goods market equilibrium satisfies the national income constraint

$$y_t = c_t + i_t + g_t + \Psi(z_t)k_{t-1}$$

Solution

We solve the model in its log-linearised form with DYNARE. The errors named in the text are given as follows

$$\epsilon_t^B = \text{Preference Shock} \qquad \eta_t^n = \text{Labour Preference Shock}$$
$$\epsilon_t^i = \text{Investment Shock} \qquad \eta_t^w = \text{Wage Mark-up Shock}$$
$$\eta_t^Q = \text{Equity Shock} \qquad \epsilon_t^g = \text{Government Spending Shock}$$
$$\epsilon_t^a = \text{Productivity Shock} \qquad \bar{\pi}_t = \text{Inflation Objective Shock}$$
$$\eta_t^p = \text{Price Mark-up Shock} \qquad \eta_t^R = \text{Monetary Shock}$$

Log-linearised model

For the empirical analysis by SW the model is log-linearised around its non-stochastic steady-state. Denoting log-deviations about equilibrium by a caret ˆ, and noting that variables dated $t+1$ are rational expectations, the log-linearised model is

$$\hat{c}_t = \frac{h}{1+h}\hat{c}_{t-1} + \frac{1}{1+h}\hat{c}_{t+1} - \frac{1-h}{(1+h)\sigma_c}\left[\left(\hat{R}_t - \hat{\pi}_{t+1}\right) + \left(\hat{\varepsilon}_t^b - \hat{\varepsilon}_{t+1}^b\right)\right]$$

$$\hat{i}_t = \frac{1}{1+\beta}\hat{i}_{t-1} + \frac{\beta}{1+\beta}\hat{i}_{t+1} + \frac{\varphi}{1+\beta}\hat{Q}_t + \beta\hat{\varepsilon}_t^i - \hat{\varepsilon}_t^i$$

$$\hat{Q}_t = -\left(\hat{R}_t - \hat{\pi}_{t+1}\right) + \frac{1-\delta}{1-\delta+\bar{r}^k}\hat{Q}_{t+1} + \frac{\bar{r}^k}{1-\delta+\bar{r}^k}\hat{r}_t^k + \hat{\eta}_t^Q$$

$$\hat{\pi}_t = \frac{\nu}{1+\beta\gamma_p}\hat{\pi}_{t-1} + \frac{\beta}{1+\beta\gamma_p}\hat{\pi}_{t+1} + \frac{(1-\beta\xi_p)(1-\xi_p)}{(1+\beta\gamma_p)\xi_p}\left[\alpha\hat{r}_t^k + (1-\alpha)\hat{w}_t - \hat{\varepsilon}_t^a + \eta_t^p\right]$$

$$\hat{w}_t = \frac{1}{1+\beta}\hat{w}_{t-1} + \frac{\beta}{1+\beta}\hat{w}_{t+1} + \frac{\gamma_w}{1+\beta}\hat{\pi}_{t-1} - \frac{1+\beta\gamma_w}{1+\beta}\hat{\pi}_t + \frac{\beta}{1+\beta}\hat{\pi}_{t+1}$$
$$- \frac{(1-\beta\xi_w)(1-\xi_w)}{(1+\beta)\left[1 + \frac{(1+\lambda_w)\sigma_n}{\lambda_w}\right]\xi_w}\left[\hat{w}_t - \sigma_n\hat{N}_t - \frac{\sigma_c}{1-h}(\hat{c}_t - h\hat{c}_{t-1}) - \hat{\varepsilon}_t^n - \eta_t^w\right]$$

$$\hat{N}_t = -\hat{w}_t + (1+\Psi)\hat{r}_t^k + \hat{k}_{t-1}$$

$$\hat{y}_t = (1-\delta k_y - g_y)\hat{c}_t + \delta k_y\hat{i}_t + g_y\hat{\varepsilon}_t^g = \phi[\hat{\varepsilon}_t^g + \alpha\hat{k}_{t-1} + \alpha\psi\hat{r}_t^k + (1-\alpha)\hat{N}_t$$

$$\hat{R}_t = \rho\hat{R}_{t-1} + (1-\rho)\left[\bar{\pi}_t + r_\pi\left(\hat{\pi}_{t-1} - \bar{\pi}_t\right) + r_y\hat{y}_t\right] + r_{\Delta\pi}\left(\hat{\pi}_t - \hat{\pi}_{t-1}\right) + r_{\Delta y}\left(\hat{y}_t - \hat{y}_{t-1}\right)$$
$$- r_a\eta_t^a - r_n\eta_t^n + \eta_t^R$$

where $\varphi = I''^{-1}$, $\beta = (1-\delta+\bar{r}^k)^{-1}$, $\psi = \frac{\Psi'(1)}{\Psi''(1)}$, $\bar{\pi}_t$ is the inflation target and the equations include various parameters which are long-run average values. Thus, there are nine endogenous variables and ten independent shocks. Five of the shocks arise from technology and preferences $(\varepsilon_t^a, \varepsilon_t^i, \varepsilon_t^b, \varepsilon_t^n, \varepsilon_t^g)$ which are generated by first-order autogressive processes, three are cost-push shocks $(\eta_t^w, \eta_t^p, \eta_t^Q)$ which are *i.i.d.* shocks and two are monetary shocks $(\bar{\pi}_t, \eta_t^R)$.