Cardiff Economics Working Papers

Working Paper No. E2018/7

The small sample properties of Indirect Inference in testing and estimating DSGE models

David Meenagh, Patrick Minford, Michael Wickens and Yongdeng Xu

March 2018

ISSN 1749-6010

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

The small sample properties of indirect inference in testing and estimating DSGE models

David Meenagh (Cardiff University)* Patrick Minford (Cardiff University and CEPR)†

Michael Wickens (Cardiff University, University of York and CEPR)‡

Yongdeng Xu (Cardiff University)§

March 2018

Abstract

Indirect inference testing can be carried out with a variety of auxiliary models. Asymptotically these different models make no difference. However, the small sample properties can differ. We explore small sample power and estimation bias both with different variable combinations and descriptive models (Vector Auto Regressions, Impulse Response Functions or Moments) in the auxiliary model. We find that both power and bias are similar when the number of variables used is the same. Raising the number of variables lowers the bias but may also raise the power unacceptably because it lowers the chances of finding a tractable model to pass the test.

Keywords: Indirect Inference, DGSE model, Auxiliary Models, Simulated Moments Method, Impulse Response Functions, VAR, Moments, power, bias

JEL Classification: C12; C32; C52; E1

1 Introduction

This paper joins a large and rapidly expanding literature on the estimation and evaluation of dynamic stochastic general equilibrium (DSGE) models by indirect inference (II). Indirect inference involves representing the data simulated from a DSGE model by an auxiliary model and comparing estimates of this auxiliary model with those obtained from observed data. The key issues are how to choose this auxiliary model and what features of the auxiliary model to exploit. In estimation the aim is to select the auxiliary model and its features to minimize bias and in testing to maximize power. In large samples these choices are often unimportant. The aim in this paper is to examine whether this is true in small samples.

The natural choice of auxiliary model for a DSGE model is a VAR. This is because, when (log-) linearized, the solution to a DSGE model can be represented by a VAR. And if any exogenous variables in the DSGE model are also represented by a VAR then the whole data set has a VAR representation (see Wickens 2014). The coefficients of the VAR solution are functions of the structural parameters. In some simple cases the precise function can be established analytically, but in general this is not possible.

An issue of concern in this paper is how the variables in the auxiliary VAR should be chosen. They could include all of the variables in the DSGE model or only a sub-set of them. Another issue is which features of the auxiliary model to exploit. We refer to these two issues as the choice of data descriptors. Possible choices of features of the auxiliary model to use are the VAR coefficients (e.g. Le et al. 2011), the impulse response functions — IRFs — (e.g. Hall et al. 2012), the Moments or the Scores. All have been used in the literature and all should give the same results asymptotically. We will focus on the small sample properties (estimation bias and test power) from using the coefficients, the IRFs, and the moments. Our results indicate that increasing the dimension of the data vector yields better small sample properties (in terms of smaller bias in estimation and large power in testing).

Estimating a DSGE model by II was originally proposed by Smith (1993), later Gourieroux et al. (1993), and the literature has expanded rapidly recently. Among them, Dridi, Guay and Renault (2007) derive

*Meenaghd@cf.ac.uk; Cardiff Business School, Cardiff University, CF10 3EU, UK

†Patrick.minford@btinternet.com; Cardiff Business School, Cardiff University, CF10 3EU, UK

‡Cardiff Business School, Cardiff University, UK

§Corresponding author; xuy16@cf.ac.uk; Cardiff Business School, Cardiff University, CF10 3EU, UK

an asymptotic distribution of the II estimator and show that it is asymptotically normal under regular conditions. Hall et al. (2012) use the impulse response function as a particular data descriptor in the auxiliary model and discuss both the small and large sample property of the II estimator. They show by Monte Carlo simulation that II estimation has good small sample properties (in terms of bias and efficiency of the estimator). These results are confirmed by Guerron-Quintana et al. (2017).

Another particularly interesting question for policy makers is how to test an already calibrated (or estimated) DSGE model by II. The II test has been discussed, for example, in Boivin and Giannoni (2006) and Dridi, Guay and Renault (2007), which mainly focus on asymptotic theory. More intensive studies have been carried out by Le et al. (2011, 2016a and b), which focus on the small sample properties of the II test. Le et al. (2011, 2016a) — using a Wald test based on the VAR coefficients - and the papers they refer to have been able to discover by Monte Carlo experiments both the power and the estimation bias for various types of structural model when different auxiliary models are chosen. In particular, they compare the II test with other model evaluation methods (e.g. Likelihood Ratio (LR) and the out of sample forecasting test) and find that the II test has much more power than the LR test (which is commonly used in evaluating a model). They give a explanation of why II has much larger power than LR test. The test could also be based on other features of the model, such as the associated impulse response functions or the moments.

Dridi, Guay and Renault (2007) propose a two-step procedure to achieve both estimation and the evaluation of misspecified DSGE models. In the first step the model is estimated using a well chosen set of moments; in the second step, the model is evaluated with chosen features of the data that the model tries to replicate. Their procedure differs in practice from those used by Guerron-Quintana et al. (2017) and Le et al. (2011, 2016a), solely in the choice of weighting matrix for the Wald statistic: Dridi, Guay and Renault (2007) use the matrix derived from the asymptotic distribution of VAR coefficients from the data-based VAR whereas Guerron-Quintana et al. (2017) and Le et al. (2011, 2016a) use the empirical distribution from the model-simulated data VARs. Le et al. (2016a) found that the use of the asymptotic weighting matrix results in a loss of power compared with the use of a simulation-based weighting matrix.

Although there is an extensive literature on the asymptotic theory of indirect inference, there is a limited literature on its small sample properties. In particular, the small sample properties of using different descriptors has not been examined in any detail.

As already noted, the use of a VAR as the auxiliary model follows naturally from the solution to a DSGE model having a VAR representation. Instead of matching the moments of the data to simulated second moments of the calibrated model, as was done in the early analysis of RBC models, matching the coefficients of an auxiliary VAR estimated on data simulated from the structural model and from the original data, or matching the corresponding impulse response functions, provides a richer set of desciptors. In their empirical analysis, Hall et al. (2012) and Guerron-Quintana et al. (2017) proposed to use the IRF as the data descriptor, while Le et al. (2011, 2016a) proposed instead the use of the VAR coefficients (and the VAR error variances). Based on asymptotic distribution theory, the three types of comparison, as they simply involve different functions of all the structural parameters, would give the same result. The interesting question that remains is whether the different data descriptors have different small sample properties.

A second issue relating to the choice of descriptors is whether the selection of variables in the auxiliary VAR affects the small-sample properties of II inference. The DSGE models that have been most often used in empirical analysis are typically similar to a Smets and Wouters (2007) type of medium-scale New Keynesian model (see, e.g. Christiano et al., 2005; Altig et al., 2011; Guerron-Quintana et al. 2013). There are normally many variables in such a model, and hence in its VAR representation. Le et al. (2011, 2016a) have found that the more of these variables that are included in the auxiliary model, the greater is the power of the test in small samples, but the lower are the chances of finding a tractable model that can pass the test. This suggests that there is a trade-off between power and finding a tractable (i.e. unrejected) model. They therefore suggest using a low-order VAR in only a few variables when using indirect inference for both the estimation of and tests of a DSGE model. This leaves open the question of which variables to include in the auxiliary model and whether this choice affects the outcome.

These questions have been raised with us by researchers and policy-makers who are users of indirect inference in model-building. In this paper we provide some new evidence on these issues and issues related to them. Minford, Wickens and Xu (2016) find that the power of the II test based on three different data descriptors (VAR coefficients, IRFs, sample moments) are similar. We extend this study of different data descriptors to the issue of estimation bias; and we also ask how both test power and estimation bias are affected by the choice of different features of the auxiliary model and different variable combinations. We do so by Monte Carlo analysis of a widely-used macro DSGE model, the Smets and Wouters (2007) model of the US.

Our main findings are that, just as the power of the II test is similar across VAR coefficients, IRFs and moments, so also small sample bias in estimation is similar and very low. We find that the choice of different variable combinations in the auxiliary model makes little difference in both the estimation and the tests. In particular, we obtain similar results whether we use any three of the variables in the model or the three principal components of the complete set of variables in the model. When more variables are included in the auxiliary model, the estimation bias falls slightly and the power of the test rises sharply, which suggests that for the Smets-Wouters model and similar macro models increasing the number of variables in an auxiliary VAR is likely to reduce the likelihood of finding a tractable model that can pass the test.

2 Indirect Inference on a DSGE model

DSGE models (possibly after linearization) have the general form:

$$\begin{aligned} A_0 E_t y_{t+1} &= A_1 y_t + B z_t \\ z_t &= R z_{t-1} + \varepsilon_t \end{aligned} \tag{1}$$

where y_t contains the endogenous variables and z_t the exogenous variables. The exogenous variables may be observable or unobservable. For example, they may be structural disturbances. We assume that z_t may be represented by an autoregressive process with disturbances ε_t that are $NID(0, \Sigma)$. Assuming that the conditions of Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson (2007) are satisfied, the solution to this model can be represented by a VAR of form

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = F \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + G \begin{bmatrix} \xi_t \\ \varepsilon_t \end{bmatrix}. \tag{2}$$

where ξ_t are innovations.

A special case of the DSGE model is where all of the exogenous variables are unobservable and may be regarded as structural shocks. An example is the Smets and Wouters (2007) US model to be examined below. This case, and its solution, can be represented as above for the complete DSGE model.

2.1 II test

The II test criterion is based on the difference between features of the auxiliary model (such as coefficients estimates, impulse response functions, moments or scores) obtained using data simulated from an estimated (or calibrated) DSGE model and those obtained using actual data; these differences are then represented by a Wald statistic, hence we call it an IIW (Indirect Inference Wald) test. The specification of the auxiliary model reflects the choice of descriptor variables.

We begin with the philosophical question of whether DSGE models can be treated as 'true' for testing purposes; in recent years a number of econometricians have dismissed DSGE models as so 'unrealistic' that one must consider them as inherently 'mis-specified'. We regard this as essentially a philosophical misunderstanding in the sense that models are not intended by construction to be 'realistic' but rather to embody economic decision-making in a logical set-up which could capture sufficient elements of economic behaviour to pass empirical tests. Such models could be termed 'pseudo-true', i.e. not close representations of 'reality' but rather abstract approximations designed to match the key data behaviour according to frequentist tests — much in the sense of Friedman (1953). Having renamed models in this way, we can proceed to test them in the usual manner. Thus the Indirect Inference test we propose supposes for the purposes of statistical testing that the model is true. Even if DSGE models are inherently 'mis-specified', it is still of interest to evaluate how misspecified they are. This test provides a very powerful way to do so. Further discussion of these philosophical issues can be found in Meenagh et al. (2018).

If the DSGE model is correct (the null hypothesis) then, whatever the descriptors chosen, the features of the auxiliary model on which the test is based will not be significantly different whether based on simulated or actual data. The simulated data from the DSGE model are obtained by bootstrapping the model using the structural shocks implied by the given (or previously estimated) model and computed from the historical data. We estimate the auxiliary model — a VAR(1) — using both the actual data and the N samples of bootstrapped data to obtain estimates a_T and $a_S(\theta_0)$ of the vector α. We then use a Wald statistic based on the difference between a_T, the estimates of the data descriptors derived from actual data, and $\overline{a_S(\theta_0)}$, the mean of their distribution based on the simulated data, which is given by:

$$WS = (a_T - \overline{a_S(\theta_0)})' W^{-1}(\theta_0)(a_T - \overline{a_S(\theta_0)})$$

where θ_0 is the vector of parameters of the DSGE model on the null hypothesis that it is true and $W(\theta_0)$ is the weighting matrix. Following Guerron-Quintana et al. (2017) and Le et al. (2011, 2016a), $W(\theta_0)$ can be obtained from the variance-covariance matrix of the distribution of simulated estimates a_S

$$W(\theta_0) = \frac{1}{N}\Sigma_{s=1}^{N}(a_s - \overline{a_s})'(a_s - \overline{a_s}) \tag{3}$$

where $\overline{a_s} = \frac{1}{N}\Sigma_{s=1}^{N}a_s$. WS is asymptotically a $\chi^2(r)$ distribution, with the number of restrictions equal to the number of elements in a_T. Appendix A shows the steps involved in finding the Wald statistic. A detailed description of the IIW test can also be found in Le et al. (2016a).

We know from Le et al. (2017) that the particular DSGE models we are examining are over-identified, so that the addition of more VAR coefficients (e.g. by raising the order of the VAR) increases the power of the test, because more nonlinear combinations of the DSGE structural coefficients need to be matched. Since these combinations are different, the chances of false values being accepted by the same samples that accept the previous combinations are low; hence the joint probability of acceptance will fall for a given degree of falseness, thereby raising the frequency of joint rejection and so the power. Le et al. (2016a) noted that increasing the power in this way also reduced the chances of finding a tractable model that would pass the test, so that there was a trade-off for users between power and tractability. The question addressed in this paper is whether, and for the same reason, adding discriptor variables also raises the power of the test.[1]

2.2 II Estimation

Estimation based on indirect inference focuses on extracting estimates of the structural parameters from estimates of the coefficients of the auxiliary model by choosing parameter values that minimise the distance between estimates of the auxiliary model based on simulated and actual data. A scalar measure of the distance may be obtained using a Wald statistic. This can be minimised using any suitable algorithm. Details of the algorithm are given in Appendix A.

The II estimation may be expressed as

$$\hat{\theta} = argmin\ WS(\theta) \tag{4}$$

Under the null hypothesis of full encompassing and some regularity conditions, Dridi, Guay and Renault (2007) show the asymptotic normality of II estimator $\widehat{\boldsymbol{\theta}}$,

$$\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \backsim N(\mathbf{0}, \Xi(N, W)) \tag{5}$$

with

$$\Xi(N,W) = \left\{\frac{\partial'(a)}{\partial(\theta_0)}W(\theta_0)^{-1}\frac{\partial'(a)}{\partial(\theta_0)}\right\}^{-1}. \tag{6}$$

$W(\theta_0)$ is the weighting matrix, which can obtained either from bootstrap samples as in equation (3):

$$W(\theta_0) = \frac{1}{N}\Sigma_{s=1}^{N}(a_s - \overline{a_s})'(a_s - \overline{a_s}) \tag{7}$$

or, as Dridi, Guay and Renault (2007) show, the optimal asymptotic weighting matrix is

$$\begin{aligned} W^*(\theta_0) \quad = \quad & J_0^{-1}I_0J_0^{-1} + \frac{1}{N}J_0^{*-1}I_0^*J_0^{*-1} + \left(1 - \frac{1}{N}\right)J_0^{*-1}K_0^*J_0^{*-1} \\ & -J_0^{-1}K_0J_0^{*-1} - J_0^{*-1}K_0'J_0^{*-1} \end{aligned} \tag{8}$$

where I_0 and J_0 is information and Hessian matrix from observed data. I_0^* and J_0^* is information and Hessian matrix from simulated data. K_0 is the covariance matrix of the score vector from two independent simulators. The score vector and Hessian matrix from observed and simulated data can be computed under a standard MLE framework. In the case where the structural model is well specified, $W(\theta_0)$ reduces to $\left(1 + \frac{1}{N}\right)J_0^{-1}(I_0 - K_0)J_0^{-1}$, since then $K_0 = K_0^*$. Also see Appendix B for the details of the asymptotic properties.

[1]Hence the decision to use a low order VAR even though the full VAR reduced form of the DSGE model will in general be of more variables and of higher order (Fernandez-Villaverde et al. 2007).

2.3 The Auxiliary Models

Different data descriptors

We consider three different features of the auxiliary model on which to assess II estimation and power: VAR coefficients, IRFs and Moments. Minford, Wickens and Xu (2016) have compared them and we briefly review their findings below. The IRFs and the Moments can be derived from the VAR coefficients and are therefore closely related (see Appendix C, repeated for convenience from Minford, Wickens and Xu, 2016); asymptotically, the three features should give the same results, but it is not clear whether this carries over in small samples.

Different data combinations

For a simple three equation New Keynesian model (Clarida, et al,. 1999), all variables can be included in the auxiliary model. However, for a medium to large scale New Keynesian DSGE model (see, e.g. Christiano et al. 2005; Smets and Wouters, 2007; Altig et al. 2011; Guerron-Quintana et al. 2013), which has been extensively discussed in the macroeconomics literature, there are many variables of interest in the model. For example, there are 14 variables (7 observed and 7 unobserved variables) in Smets-Wouters 2007 and in the Guerron-Quintana et al. 2017 model. A practical question is to consider which variable combinations should be used in VAR model estimation. In II estimation, one may consider including all variables in the VAR model (Guerron-Quintana et al. 2017). However, there is a cost in terms of the computation time which increases significantly. In the II test, Le et al. (2011, 2016a) find that the more variables that are included, the higher is the power of the II test. This is because there are more features of the data that need to be jointly matched, each of which is, in general, a nonlinear combination of the structural parameters. Consequently, the lower is the chance of finding a tractable model that can pass the test. In other words, there is a trade-off between power and tractability. The authors suggest using a three variable VAR(1) as the auxiliary model. Again an interesting question is whether there would be any difference if one included a different three variable combination in the VAR model.

A further alternative is to use the first three principal components of a wider set of possible descriptors. This may be better than arbitrarily choosing three variables from the set of potential descriptors. This involves incorporating a standard factor analysis within the II process: thus the data, whether from the actual sample or from the samples generated by simulation, has the three factors extracted from it which are then used in the subsequent VAR estimation.

3 Monte Carlo Experiments

We use Monte Carlo simulation to explore the two issues raised above based on a popular model in the macroeconomics literature: the Smets and Wouters model (2007) DSGE model. A sample size of 200 is chosen, as this is typical for macro data. We compare the power of the II test and the II estimates using different data descriptors and different variable combinations. Our aim is to see whether better results in terms of higher test power or lower small sample estimation bias can be found by altering the choice of descriptors or variables.

3.1 Comparing the Power of the II test

We follow the same approach as Le et al. (2016a). Specifically, we generate falseness by introducing a rising degree of numerical mis-specification for the model parameters. Thus we construct a model whose parameters are moved $x\%$ away from their true values in both directions ($+/-$ alternatively); similarly the higher moments of the error processes (standard deviation) are altered by the same $x\%$.[2] For all the experiments, the eigenvalues of reduced form VAR coefficients are all strictly less than unity in modulus, so Fernandez-Villaverde et al.'s (2007) condition that the DSGE model has a VAR representation is satisfied. We create 1000 samples from the true model: then we obtain from these samples the distribution of the Wald statistic by bootstrapping the false model (the bootstrap number is 500) as if it is true. We use this distribution to assess how many times the $x\%$ False model is rejected with 95% confidence; notice that this fixes the size of the test throughout at 5%. The Monte Carlo simulation results are presented in Tables 2 and 3, where y: real GDP, pi : inflation rate, r : real interest rate, c : consumption, i : investment, l : employment, q : Tobin's q, w : the real wage.

[2] See Le et al. (2016b) section 4.1 for full details of the experiments.

1) Comparing the Power of the II test across data descriptors

Table 1: Power of the II test across data descriptors

	Rejection Rate							
Degree of falseness	0	1%	3%	5%	7%	10%	15%	20%
VAR coefficients	0.05	0.128	0.866	0.997	1.000	1.000	1.000	1.000
IRF	0.05	0.140	0.852	0.998	1.000	1.000	1.000	1.000
Moment	0.05	0.114	0.326	0.665	0.913	0.997	1.000	1.000

Notes: Three variables are used in VAR are (y, pi, r), as in Le et al. (2011)
Source: Minford, Wickens and Xu (2016)

We reproduce these results from Minford, Wickens and Xu (2016) for completeness; details can be found there. What is rather remarkable about these comparisons is how similar the power is across all three methods for the VAR coefficients and IRF as data descriptors.

2) Comparing the Power of the II test across variable choice

Le et al. (2011, 2016a, 2017) use the variable combination (y, pi, r) in their II test. In this experiment, we consider in addition two more variable combinations: (c, i, l) and (q, w, r). The results are reported in Table 2.

Table 2: Power of the II test across variable choice (3 variables)

	Rejection Rate							
Degree of falseness	0	1%	3%	5%	7%	10%	15%	20%
(y, pi, r)	0.05	0.128	0.866	0.997	1.000	1.000	1.000	1.000
(c, i, l)	0.05	0.094	0.561	0.923	0.986	1.000	1.000	1.000
(q, w, r)	0.05	0.072	0.276	0.771	0.984	1.000	1.000	1.000

Notes: VAR coefficients are used as data descriptors, as in Le et al. (2011)

The combinations all produce tests with very similar power.

Next we consider groups of four variables by adding a variable to the previous combinations. The results are reported in Table 3. As we can see, this raises power consistently for each group. But the power, while similar for each group, preserves much the same differences as in the three-variable sets.

Table 3: Power of the II test across variable choice (4 variables)

	Rejection Rate							
Degree of falseness	0	1%	3%	5%	7%	10%	15%	20%
(y, pi, r, i)	0.05	0.166	1.000	1.000	1.000	1.000	1.000	1.000
(c, i, l, r)	0.05	0.131	0.896	0.977	1.000	1.000	1.000	1.000
(q, w, r, c)	0.05	0.118	0.658	0.991	1.000	1.000	1.000	1.000

Notes: VAR coefficients are used as data descriptors, as in Le et al. (2011)

In all cases these four-variable sets give huge power, so that even 3% falsity induces almost 100% rejection. A test with such high power would make it likely that only a model extremely close to reality would not be rejected. In practice, users may therefore prefer to use the less powerful 3-variable sets.

Table 4 reports the power of the test when using all seven variables and when using the three principal components of the seven variables. Not surprisingly, using seven variables gives a test with even larger power than one with the three or four variables. 1% falsity induces 37% rejection and 3% falsity induces 100% rejection whereas the three principal components have similar power to the previous four variable combinations.

To summarise, including more descriptors improves the power of the test, but if one wishes to limit the number of descriptors, any three variables gives similar power. This includes the use of principal components, which effectively coincide with the use of output, inflation and interest rates as the three variables chosen, as indeed these three factors are likely to largely coincide with these three.

Table 4: Power of the II test across variable choice (7 variables)

	Rejection Rate							
Degree of falseness	0	1%	3%	5%	7%	10%	15%	20%
All 7 Variables	0.05	0.368	1.000	1.000	1.000	1.000	1.000	1.000
3 PCs	0.05	0.142	0.843	0.998	1.000	1.000	1.000	1.000

Notes: VAR coefficients are used as data descriptors, as in Le et al. (2011)

3.2 Comparing II estimation in small samples

In the following subsection, we examine the effect in small samples of the choices of different data descriptors and different variable combinations on the estimation bias of the parameters of the Smets-Wouters model. The True values are the estimates reported by Smets-Wouters.

1) Comparing II estimation across data descriptors

While we know from earlier work that the estimators have similar asymptotic properties, there is no work comparing their small sample properties. We use the same Monte Carlo experiment with the Smets and Wouters (2007) model treated as the true model; and we re-estimate the model for each of 100 samples of data simulated from the true specification of the model. The true parameter values are from Smets and Wouters (2007), table 4. Also see Appendix D for the parameter descriptions.

In estimation, we start the initial parameter values by falsifying them by 10% in both directions (+/− alternately). We then estimate each sample and report the Mean and average bias of the II estimators. The results are reported in the Table 5.

Table 5: Bias of II estimates by using different data descriptors

Data Descriptor		VAR Coefficients			IRF			Moments		
		II estimates			II estimates			II estimates		
True Values		Mean	Std	Bias	Mean	Std	Bias	Mean	Std	Bias
α	0.19	0.189	0.020	0.32%	0.191	0.021	0.37%	0.197	0.018	1.50%
h	0.71	0.692	0.065	2.51%	0.684	0.069	3.61%	0.669	0.049	2.77%
ι_p	0.22	0.221	0.023	0.45%	0.223	0.026	1.32%	0.229	0.021	4.08%
ι_w	0.59	0.586	0.060	0.69%	0.584	0.068	0.96%	0.570	0.058	3.40%
ξ_p	0.65	0.664	0.068	2.08%	0.659	0.074	1.35%	0.674	0.063	3.71%
ξ_w	0.73	0.725	0.075	0.71%	0.731	0.083	0.17%	0.702	0.073	3.82%
φ	5.48	5.531	0.557	0.93%	5.555	0.608	1.36%	5.712	0.524	4.24%
Φ	1.61	1.607	0.166	0.16%	1.635	0.182	1.53%	1.576	0.140	2.14%
ψ	0.54	0.540	0.057	0.00%	0.543	0.062	0.57%	0.561	0.051	3.88%
$r_{\Delta y}$	0.22	0.222	0.023	1.05%	0.222	0.024	1.09%	0.210	0.020	1.38%
ρ	0.81	0.787	0.047	2.80%	0.770	0.060	4.95%	0.842	0.058	3.91%
r_π	2.03	2.007	0.188	1.12%	1.992	0.204	1.86%	1.965	0.188	3.20%
r_y	0.08	0.079	0.009	1.25%	0.081	0.009	1.38%	0.082	0.008	2.96%
σ_c	1.39	1.358	0.137	2.29%	1.377	0.155	0.95%	1.356	0.132	2.45%
σ_l	1.92	1.930	0.208	0.54%	1.944	0.211	1.24%	1.981	0.191	1.18%
Average			0.113	1.13%		0.124	1.51%		0.106	3.04%

Notes: The true parameter values are from Smets and Wouters (2007), table 4. Three variables used in VAR are (y, pi, r), as in Le et al. (2011). Std denotes the standard deviation. Bias denotes the bias of II estimates.

We find that the II estimator has very small bias. The average absolute biases of the II estimator based on using VAR coefficients and the IRFs as auxiliary models are 1.1% and 1.5% respectively. This very low bias is related to the high power of the IIW test which, as we have seen, rejects parameters as little as 7% distant from the true values 100% of the time, regardless of which three-variable combination is used. Le et al. (2016a) find that the comparable FIML estimates are heavily biased in small samples. II estimation is, by contrast, found to be almost unbiased, which is clearly a very useful property for those using DSGE models in practice. The average bias of the II estimator based on moments as the auxiliary model is somewhat higher at 3%, but is also more efficient.

The II estimator is therefore not much affected by varying the three different data descriptors in the auxiliary model.

2) Comparing the II estimation across variable choice

Next we compare the II estimates in small samples across different variable combinations. We use the same sets of variable combinations as those in the II test experiment. We start the initial parameter values by falsifying them by 10% on both directions (+/− alternatively). We then estimate each sample and report the Mean and average bias. The results are reported in the Table 6.

Table 6: Bias of II estimates by using different variable combinations

Var Comb		(y, pi, r)			(c, i, l)			(q, w, r)			3 Factors for the 7 variables		
		II estimates			II estimates			II estimates			II estimates		
True Values		Mean	Std	Bias	Mean	Std	Bias	Mean	Std	Bias	Mean	Std	Bias
α	0.19	0.189	0.020	0.32%	0.188	0.021	1.2%	0.193	0.020	1.74%	0.194	0.023	1.89%
h	0.71	0.692	0.065	2.51%	0.711	0.073	0.2%	0.674	0.057	5.08%	0.692	0.081	2.55%
ι_p	0.22	0.221	0.023	0.45%	0.215	0.025	2.3%	0.223	0.024	1.40%	0.225	0.027	2.12%
ι_w	0.59	0.586	0.060	0.69%	0.571	0.068	3.2%	0.580	0.063	1.70%	0.583	0.072	1.21%
ξ_p	0.65	0.664	0.068	2.08%	0.635	0.073	2.3%	0.666	0.069	2.53%	0.652	0.082	0.28%
ξ_w	0.73	0.725	0.075	0.71%	0.727	0.082	0.4%	0.719	0.080	1.49%	0.732	0.088	0.25%
φ	5.48	5.531	0.557	0.93%	5.385	0.613	1.7%	5.575	0.579	1.73%	5.554	0.676	1.35%
Φ	1.61	1.607	0.166	0.16%	1.595	0.141	0.9%	1.583	0.172	1.69%	1.600	0.199	0.59%
ψ	0.54	0.540	0.057	0.00%	0.524	0.062	3.0%	0.553	0.057	2.36%	0.541	0.069	0.13%
$r_{\Delta y}$	0.22	0.222	0.023	1.05%	0.218	0.025	1.1%	0.217	0.024	1.49%	0.220	0.027	0.17%
ρ	0.81	0.787	0.047	2.80%	0.797	0.088	1.6%	0.816	0.065	0.74%	0.809	0.096	0.10%
r_π	2.03	2.007	0.188	1.12%	2.015	0.234	0.7%	1.983	0.205	2.32%	2.014	0.246	0.80%
r_y	0.08	0.079	0.009	1.25%	0.078	0.009	2.3%	0.080	0.009	0.10%	0.081	0.010	1.47%
σ_c	1.39	1.358	0.137	2.29%	1.370	0.156	1.4%	1.345	0.141	3.26%	1.370	0.168	1.46%
σ_l	1.92	1.930	0.208	0.54%	1.876	0.218	2.3%	1.957	0.204	1.93%	1.937	0.234	0.90%
Average			0.113	1.13%		0.123	1.70%		0.118	1.97%		0.140	1.02%

Notes: The true parameter values are from Smets and Wouters (2007), table 4. VAR coefficients are used as data descriptors, as in Le et al. (2011). Std denotes the standard deviation. Bias denotes the bias of II estimates.

The biases for individual coefficients vary across the different variable combinations while the average bias across all of the coefficients is least using three principal components. Nevertheless the bias is very small in absolute size across all four sets. Using three principal components the bias is similar to the variable combination of output, inflation and interest rates. This set is 'central' in the sense that it summarises the key dimensions of macro variability for real variables, nominal variables, and yields. Clearly, the three principal components are summarising the effects of these variables. With three factors there is a loss of efficiency, with the standard deviation rising by about a quarter. This could be due to the large rise (100% or more) in the computational burden due to extracting the principal components.

Raising each group to four variables reduces the bias of the II estimators, as shown in Table 7; the average bias is reduced to a range of $1.2 - 1.5\%$. Again this is consistent with the higher test power of four variables. Raising the number of variables to seven reduces the bias further but efficiency falls because the standard deviation of the estimates rises. Again we think the loss of efficiency may be due to the large rise in computational burden on a PC and could be eliminated by using a large machine with many more bootstraps. As with the three factor case we have not pursued this issue here, as plainly using 7 variables would create greatly excessive power for the test.

To summarise, we find that the estimation biases of the II estimator based on VAR coefficients and IRFs are similar and very low (about 1.5%), while that based on moments is only slightly higher. The biases are not much affected by the choice of descriptor variable set. In particular, using the three main variables gives very low small sample biases. There is generally a reduction in bias from including more variables in the auxiliary model, as one would expect from the resulting increase in test power.

Table 7: Bias of II estimates by using different variable combinations

Var Comb		(y, pi, r, i)			(c, i, l, r)			(q, w, r, c)			All 7 Variables		
		II estimates			II estimates			II estimates			II estimates		
True Values		Mean	Std	Bias	Mean	Std	Bias	Mean	Std	Bias	Mean	Std	Bias
α	0.19	0.196	0.019	3.32%	0.191	0.022	0.68%	0.195	0.021	1.45%	0.193	0.026	1.38%
h	0.71	0.690	0.058	2.86%	0.685	0.068	3.54%	0.682	0.063	2.02%	0.699	0.096	1.59%
ι_p	0.22	0.221	0.023	0.24%	0.222	0.025	1.05%	0.223	0.025	1.37%	0.220	0.031	0.16%
ι_w	0.59	0.582	0.063	1.32%	0.582	0.066	1.32%	0.577	0.063	2.26%	0.589	0.079	0.20%
ξ_p	0.65	0.658	0.070	1.24%	0.657	0.074	1.06%	0.669	0.068	1.95%	0.659	0.085	1.34%
ξ_w	0.73	0.721	0.076	1.21%	0.733	0.081	0.40%	0.723	0.081	1.00%	0.724	0.101	0.89%
φ	5.48	5.565	0.590	1.54%	5.540	0.621	1.10%	5.568	0.613	1.60%	5.503	0.751	0.41%
Φ	1.61	1.577	0.166	2.08%	1.595	0.183	0.92%	1.596	0.176	0.87%	1.596	0.218	0.88%
ψ	0.54	0.543	0.057	0.50%	0.550	0.060	1.76%	0.548	0.061	1.51%	0.544	0.071	0.70%
$r_{\Delta y}$	0.22	0.220	0.023	0.17%	0.226	0.023	2.64%	0.217	0.024	1.25%	0.219	0.030	0.34%
ρ	0.81	0.792	0.047	2.29%	0.780	0.054	3.67%	0.804	0.071	0.80%	0.812	0.108	0.19%
r_π	2.03	2.028	0.199	0.08%	2.005	0.201	1.22%	2.004	0.210	1.28%	2.005	0.271	1.24%
r_y	0.08	0.080	0.009	0.19%	0.082	0.009	2.00%	0.081	0.009	1.35%	0.081	0.011	1.17%
σ_c	1.39	1.387	0.137	0.19%	1.345	0.151	3.21%	1.357	0.148	2.35%	1.377	0.187	0.93%
σ_l	1.92	1.948	0.203	1.46%	1.938	0.215	0.94%	1.948	0.211	1.45%	1.967	0.260	2.45%
Average			0.116	1.24%		0.106	1.28%		0.123	1.50%		0.154	0.92%

Notes: The true parameter values are from Smets and Wouters (2007), table 4. VAR coefficients are used as data descriptors, as in Le et al. (2011). Std denotes the standard deviation. Bias denotes the bias of II estimates.

4 Evaluating the Smets and Wouters model in practice

In a final exercise we re-estimated the DSGE model of Le, Meenagh and Minford (2016b). This was originally estimated on the first three-variable set. We have re-estimated it on investment, the real wage and the interest rate and found a set of coefficients that passed the II test with a p-value of 0.0639. The estimated parameters differed in absolute value on average by only 4.1%, and the policy results under different monetary policies were almost identical. Table 8 shows that the stability of the different policies, as measured by the frequency of crises, are very similar under both sets of coefficients. This example confirms our general finding that the choice of variable set has little effect on the results of tests or estimation based on indirect inference.

Table 8: Stability under different monetary regimes

Monetary policy	Base Case			NGDPT
	y, π, r	inv, w, r	y, π, r	inv, w, r
Frequency of crisis per 1000 years	20.8	21.4	1.83	2.04

Note: Base case denotes Monetary Base rule (responds to credit premium). NGDPT denotes Nominal GDP target (NGDPT) for inflation.

5 Conclusions

We set out in this paper to answer a frequently asked question: does it make any difference in Indirect Inference which features of an auxiliary model and which set of variables are chosen to be the data descriptors to be matched by the structural model under investigation? Based on Monte Carlo experiments with the widely-used Smets and Wouters (2007) model of the US we find that the power of II tests rises dramatically as the number of variables increases, making it very difficult or impossible to find a tractable model capable of passing the test. Whether basing II tests or estimation on the VAR coefficients, the IRFs, or on moments a small set of variables — typically three — or the same number of principal components, appears to be sufficient to achieve high power and small biases.

Our results also show that there is hardly any difference either in power or estimation bias between different variable sets, including the three main factors driving the data. All three variable sets have such similar properties that it seems unlikely users will notice much difference in practice. This is confirmed when we perform a similar analysis on the version of the model examined in Le et al. (2016a).

As expected, when the number of variables included rises to four, all sets have greatly enhanced power and much-reduced estimation bias; however, in practice, this power makes it unlikely that a tractable model can be found to pass the test; using fewer may therefore be preferable.

These findings appear to reflect the general equilibrium character of a macroeconomy and structural models of it. The solutions to these models show that the behaviour of the variables determined within the system have strongly related dynamics which are driven by a common set of parameters and exogenous shocks and their data behaviour will be driven by reduced forms that consist of similar numbers of different nonlinear combinations of the structural parameters. If more of these are added to the test, the test is more demanding with higher power and the estimation bias falls but the particular features of the data that are commonly used to carry out the estimation and tests does not much matter.

The II test therefore seems to be effective in producing tests with high power and estimates with low bias even when based on highly selective information and small samples. This econometric technology appears to open up wide possibilities for empirical macroeconomics, an area that is at present largely neglected.

References

[1] Altig, D., Christiano, L. J., Eichenbaum, M., & Linde, J. , 2011. Firm-specific capital, nominal rigidities and the business cycle. Review of Economic Dynamics, 14(2), 225-247.

[2] Boivin, J. and Giannoni, M. P., 2006. Has Monetary Policy Become More Effective?. Review of Economics and Statistics, 88(3), 445-462.

[3] Christiano, L.J., M. Eichenbaum and C. Evans, 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. Journal of Political Economy, 113,1-45.

[4] Clarida, R., Gali, J. and Gertler, M. L., 1999. The Science of Monetary Policy: A New Keynesian Perspective. Journal of Economic Literature, 37(4), 1661-1707.

[5] Dridi, R., Guay, A. and Renault, E., 2007. Indirect inference and calibration of dynamic stochastic general equilibrium models. Journal of Econometrics, 136(2), 397-430.

[6] Fernandez-Villaverde, J., Rubio-Ramirez, F., Sargent, T, and Watson, M. ,2007. ABCs (and Ds) of Understanding VARs. American Economic Review, 1021-1026.

[7] Friedman, M. (1953). "The Methodology of Positive Economics".." In Essays in Positive Economics. Chicago: University of Chicago Press: 3–43.

[8] Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference, Journal of Applied Econometrics 8, S85–S118.

[9] Guerron-Quintana, P., Inoue, A. and Kilian,L., 2013. Frequentist Inference in Weakly Identified DSGE Models. Quantitative Economics, 4, 197-229.

[10] Guerron-Quintana, P., Inoue, A. and Kilian,L., 2017. Impulse Response Matching Estimators for DSGE Models. Journal of Econometrics, 2017, 196, (1), 144-155.

[11] Hall, A.R., A. Inoue, J.M. Nason, and Rossi, B., 2012. Information Criteria for Impulse Response Function Matching Estimation of DSGE Models. Journal of Econometrics, 170, 499-518.

[12] Le, V.P.M., Meenagh, D., Minford, P., Wickens, M., 2011. How much nominal rigidity is there in the US economy — testing a New Keynesian model using indirect inference. Journal of Economic Dynamics and Control 35(12), 2078–2104.

[13] Le, V.P.M., Meenagh, D., Minford, P.,Wickens, M., and Xu, Y., 2016a. Testing macro models by indirect inference: a survey for users. Open Economies Review 27, 1-38.

[14] Le, V.P.M., Meenagh, D., Minford, P., 2016b. Monetarism rides again? US monetary policy in a world of Quantitative Easing. Journal of International Financial Markets Institutions and Money, 44, (C), 85-102

[15] Le, V.P.M, Meenagh, D., Minford, P., Wickens, 2017. A Monte Carlo procedure for checking identification in DSGE models. Journal of Economic Dynamics and Control, 2017, 76, (C), 202-210

[16] Meenagh, D., Minford, P.,Wickens, M., and Xu, Y., 2018. Testing DSGE models by indirect inference: a survey of recent findings. Cardiff Economics Working Papers E2018/14, Cardiff University, Cardiff Business School, Economics Section.

[17] Minford, P., Wickens, M., and Xu, Y, 2016. Comparing different data descriptors in Indirect Inference tests on DSGE models, Economics Letters, 2016, 145, (C), 157-161

[18] Smets, F., Wouters, R., 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. American Economic Review, 97, 586–606.

[19] Smith, A., 1993. Estimating nonlinear time-series models using simulated vector autoregressions, Journal of Applied Econometrics 8, S63–S84.

[20] Wickens, M.R., 2014. How useful are DSGE macroeconomic models for forecasting? Open Economies Review 25, 171-193.

Appendix A: Deriving the Wald statistic

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 θ_0*.*

Estimate the structural errors ε_t of the DSGE macroeconomic model, $x_t(\theta_0)$, given the stated values θ_0 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. Where the equations contain no expectations the errors can simply be backed out of the equation and the data. Where there are expectations, estimation is required for the expectations; here we carry this out using the robust instrumental variables methods of McCallum (1976) and Wickens (1982), with the lagged endogenous data as instruments — thus effectively we use the auxiliary model VAR. An alternative method for expectations estimation is the "exact" method; here we use the model itself to project the expectations and because these depend on the extracted residuals there is iteration between the two elements until convergence.

Step 2: Derive the simulated data

Under the null hypothesis the $\{\varepsilon_t\}_{t=1}^{T}$ are the structural errors. The simulated disturbances are drawn from these errors. In some DSGE models, including the Smets-Wouters model, many of the structural errors are assumed to be generated by autoregressive processes rather than being serially independent. If they are, then under our method we need to estimate them. We derive the simulated data by drawing the bootstrapped disturbances by time vector to preserve any simultaneity between them, and solving the resulting model using Dynare (Juillard, 2001). To obtain the N bootstrapped simulations we repeat this, drawing each sample independently.

Step 3: Compute the Wald statistic

We estimate the auxiliary model — a VAR(1) — using both the actual data and the N samples of simulated data to obtain estimates a_T and $a_S(\theta_0)$ of the vector a. The distribution of $a_T - \overline{a_S(\theta_0)}$ and its covariance matrix $W(\theta_0)^{-1}$ are estimated by bootstrapping $a_S(\theta_0)$. The bootstrapping proceeds by drawing N bootstrap samples of the structural model, and estimating the auxiliary VAR on each, thus obtaining N values of $a_S(\theta_0)$; we obtain the covariance of the simulated variables directly from the bootstrap samples. The resulting set of a_k vectors ($k = 1, ..., N$) represents the sampling variation implied by the structural model from which estimates of its mean, covariance matrix and confidence bounds may be calculated directly. Thus, the estimate of $W(\theta_0)$ is

$$W(\theta_0) = \frac{1}{N}\Sigma_{k=1}^{N}(a_k - \overline{a_k})'(a_k - \overline{a_k}) \qquad (9)$$

where $\overline{a_k} = \frac{1}{N}\Sigma_{k=1}^{N}a_k$. We then calculate the Wald statistic for the data sample; we estimate the bootstrap distribution of the Wald from the N bootstrap samples. The Wald statistics are given by

$$WS = (a_T - \bar{a}_s(\theta_o))'W(a_s(\theta_o))^{-1}(a_T - \bar{a}_s(\theta_o)) \qquad (10)$$

We note that the auxiliary model used is a VAR(1) and is for a limited number of key variables. By raising the lag order of the VAR and increasing the number of variables, the stringency of the overall test of the model is increased. If 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.

Appendix B: The asymptotic distribution

Suppose the actual data consist in the observation of a stochastic process $\{y_t\}_{t=1}^T$ or $\{Y^0, X\}$. Then for each given value of the parameters θ_0 in the structural model, it is possible to simulate data $\{y_t^s\}_{t=1}^T$ or $\{Y^{0s}, X^s\}$ conditional on the observed data and for given initial conditions. This is done by the bootstrap process discussed in Appendix A.

To derive the asymptotic theory of indirect inference, we need the information and Hessian matrix from observed and simulated data.These may differ, as shown in Dridi, Guay and Renault (2007). We must also consider a second set of similar matrices associated with the simulator when the pseudo-true values of the parameters are used for simulation. More precisely, we define:

$$\begin{aligned} I_0^* &= \frac{1}{N}\sum_{s=1}^{N} E(S_t^s(Y^{0s}, X^s)S_t^s(Y^{0s}, X^s)') \\ J_0^* &= -\frac{1}{N}\sum_{s=1}^{N} E(H_t^s(Y^{0s}, X^s)) \\ K_0 &= \frac{1}{N}\sum_{s=1}^{N} E(S_t(Y^0, X)S_t^*(Y^{0*}, X^*)) \\ K_0^* &= \frac{1}{N(N-1)/2}\sum_{s\neq l}^{N} E(S_t^s(Y^{0s}, X^s)S_t^l(Y^{0l}, X^l)) \end{aligned} \tag{11}$$

where $S_t(.)$ is the score vector and $H_t(.)$ is the Hessian matrix. The score vector and Hessian matrix from observed and simulated data can be computed under a standard MLE framework. K_0^* is the covariance matrix of the score vector from two independent simulators $\{Y^{0s}, X^s\}$ and (Y^{0l}, X^l) for $s \neq l$.

Under the null hypothesis of full encompassing and some regularity conditions, Dridi, Guay and Renault (2007) show that the distribution of the II estimator $\widehat{\boldsymbol{\theta}}$ is asymptotic normal

$$\sqrt{T}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \backsim N(\mathbf{0}, \Xi(N, W) \tag{12}$$

with

$$\Xi(N, W) = \left\{\frac{\partial'(\boldsymbol{\delta})}{\partial(\theta_0)} W(\theta_0)^{-1}\frac{\partial'(\boldsymbol{\delta})}{\partial(\theta_0)}\right\}^{-1}. \tag{13}$$

and an asymptotic weighting matrix

$$\begin{aligned} W(\theta_0) &= J_0^{-1}I_0J_0^{-1} + \frac{1}{N}J_0^{*-1}I_0^*J_0^{*-1} + \left(1 - \frac{1}{N}\right)J_0^{*-1}K_0^*J_0^{*-1} \\ &\quad - J_0^{-1}K_0J_0^{*-1} - J_0^{*-1}K_0'J_0^{*-1} \end{aligned} \tag{14}$$

When the structural model is well specified $K_0 = K_0^*$ and $W(\theta_0)$ reduces to $\left(1 + \frac{1}{N}\right)J_0^{-1}(I_0 - K_0)J_0^{-1}$.

Appendix C: The three data descriptors (source: Minford, Wickens and Xu, 2016)

For simplicity, we use first order VAR, as an auxiliary model, to compare the three data descriptors,

$$y_t = A_1 y_{t-1} + u_t \tag{15}$$

where u_t is assumed to be $NID(0, \Sigma)$. The MLE/OLS estimates of $\hat{A}_1$ and $\hat{\Sigma}$ are:

$$\hat{A}_1 = (\sum_{t=2}^{T} y_{t-1}y_{t-1}')^{-1}\sum_{t=2}^{T} y_{t-1}y_t'$$

$$\hat{\Sigma} = \frac{\sum_{t=2}^{T}\hat{u}_t\hat{u}_t'}{n-k} = \frac{\sum_{t=2}^{T}(y_{t-1} - \hat{A}_1y_{t-1})(y_{t-1} - \hat{A}_1y_{t-1})'}{n-k} \tag{16}$$

The error u_t is related to the structural innovations of the DSGE model ξ_t as $u_t{=}B\xi_t$, where ξ_{it} is uncorrelated with ξ_{jt} for $i \neq j$. We assume B is known (and imposed from the model being tested) so that we can identify the structural errors causing the impulses. The IRF to the shocks of the structural errors is then:

$$IRF(h) = \frac{dy_{t+h}}{du_t'} = {A_1}^{(h-1)}B, \qquad h = 0, 1, 2, ... \tag{17}$$

The average of IRF over M periods is defined as

$$IRF_{Ave} = \frac{1}{M+1}\sum_{h=0}^{M} IRF(h) \qquad . \tag{18}$$

The asymptotic second moments of the y_t process can be derived as:

$$\begin{aligned} \Gamma_y(h) &= E(y_t - \mu)(y_{t-h} - \mu)^{'} \\ &= \lim_{n\to\propto} \sum_{i=0}^{n}\sum_{j=0}^{n} A_1^i E(u_{t-i}u_{t-h-i}')A_1^{j}{}' \\ &= \sum_{i=0}^{\propto} A_1^{h+i}\Sigma A_1^{i}{}' \end{aligned} \tag{19}$$

as $E(u_t u_s^{'}) = 0$ for $s \neq t$ and $E(u_t u_t^{'})$ $=\Sigma$ for t.

The covariance matrix can be obtained by setting $h = 0$,

$$\begin{aligned} \Gamma_0 &= E(y_t - \mu)(y_t - \mu)' \\ &= \sum_{i=0}^{\propto} A_1^i \Sigma A_1^{i}{}' \end{aligned} \tag{20}$$

When one compares the IIW statistics, one finds:

- With VAR coefficients as the data descriptors, the II test/estimation uses the estimated VAR coefficients, as given in eq (16).
- With IRF functions as the descriptors, the II test/estimation uses the estimated IRF functions, as given in eq (17), which reveals that the IRF function is a nonlinear function of VAR coefficients and the error covariance matrix (which is identified by the B matrix). If we considered the IRF over 4 years (16 periods) and took its average, then this average of IRF has 9 elements for a 3 variable VAR (1) model. This equals the number of VAR coefficients. So the test utilises a comparable number of descriptors. We here take averages of IRFs for different shock/variable combinations.
- With the Simulated Moments (SM) as the data descriptors, the test uses the simulated moments of the data. Consider the covariance matrix, and use its lower triangular elements. For a 3-variable VAR model, we have 3(3+1)/2=6 elements to compare. The first order autocorrelation coefficients are added as additional moments. This brings the number of elements in the Wald statistic again to 9. From the theoretical moments derived above, we know that the data covariance is a nonlinear combination of VAR coefficients and the error covariance matrix. Again the number of descriptors is comparable with the number of VAR descriptors.

Appendix D: Parameter Descriptions

Table 9: Variable Descriptions

Variable Name	True Value	Description
α	0.19	Income share of capital
h	0.71	External habit formation
ι_p	0.22	Degree of price indexation
ι_w	0.59	Degree of wage indexation
ξ_p	0.65	Degree of price stickiness
ξ_w	0.73	Degree of wage stickiness
φ	5.48	Elasticity of the capital adjustment cost function
Φ	1.61	1+the share of fixed costs in production
ψ	0.54	Elasticity of the capital utilization adjustment cost
$r_{\Delta y}$	0.22	Taylor rule coefficient
ρ	0.81	Taylor rule coefficient (interest rate smoothing)
r_π	2.03	Taylor rule coefficient
r_y	0.08	Taylor rule coefficient
σ_c	1.39	Elasticity of intertemporal substitution for labour
σ_l	1.92	Elasticity of labour supply to real wage

Note: The true parameter values are from Smets and Wouters (2007), table 4.