Testing part of a DSGE model by Indirect Inference*

Patrick Minford (Cardiff University and CEPR)†
Michael Wickens (Cardiff University, University of York and CEPR)‡
Yongdeng Xu (Cardiff University)§

May 2018

Abstract

We propose a way of testing a subset of equations of a DSGE model. The test draws on statistical inference for limited information models and the use of indirect inference to test DSGE models. Using the numerical small sample distribution of our test for two subsets of equations of the Smets-Wouters model we show that the test has accurate size and good power in small samples, and better power than using asymptotic distribution theory. In a test of the Smets-Wouters model on US Great Moderation data we reject the specification of the wage-price but not the expenditure sector. This points to the wage-price sector as the source of overall model rejection.

Keywords: sub sectors of models, limited information, indirect inference, testing DSGE models equations, Monte Carlo, power, test size

JEL Classification: C12; C32; C52; E1

---

*We are grateful for most helpful comments on an earlier version from the editor, Jonathan Temple, and two referees. All remaining errors are our responsibility.

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

‡Cardiff Business School, Cardiff University, CF10 3EU, UK

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

1 Introduction

In this paper we show how to test subsets of the structural equations of a DSGE model using indirect inference (II). Although a key feature of DSGE models is that they represent general equilibrium, implying that the model is a complete structural system, it may be of interest to examine individual structural equations or subsets of equations. One reason is that econometric tests of DSGE models, although rare, commonly lead to their rejection. It may therefore be useful to find which of the structural equations is causing the whole model to be rejected. A more positive reason for performing subset tests is that we can also discover which parts of the model are not rejected and so do not need respecifying.

This problem falls under the category of limited information inference. In deriving our test we make use of a key result on statistical inference for limited information models by Phillips and Wickens (1978) and Godfrey and Wickens (1982). Given a subset of the structural equations of an econometric model, they show that a simple way to estimate and test these equations is to augment them with unrestricted reduced-form equations of the endogenous variables not explained by, but included in, the subset. Full information statistical procedures can then be used on the resulting complete model. For example, using full information maximum likelihood estimation would give the standard limited information maximum likelihood estimator. This method can easily be adapted for use in indirect inference and applied to DSGE models.

The method of indirect inference was first proposed by Gourieroux, Monfort and Renault (1993), Gregory and Smith (1993) and Smith (1993). It was shown by Le et al. (2011, 2016) to be well-suited to estimating and testing DSGE models, notably those already estimated by Bayesian methods. Le et al. (2011) tested the widely-cited Smets-Wouters (2007) model of the United States using indirect inference and found it was rejected but, when respecified to have greater price flexibility, it was not rejected. These findings provide part of the motivation for the application of the subset test proposed in this paper to the wage-price sector of the Smets-Wouters model. Le et al. (2016) showed that the power of tests of DSGE models based on indirect inference was very high and better than other tests such as likelihood ratio tests. In this paper we examine whether the good power property of indirect inference tests carries over to testing a subset of a DSGE model.

The solution of a (linearized) DSGE model is a VAR in the endogenous variables of the model plus conditional expectations of its current and future exogenous variables. If the exogenous variables are represented by a VAR then solution of the complete set of variables - endogenous and exogenous - can be represented by a VAR involving both sets of variables, see Wickens (2014).[1] The coefficients of this VAR model of the full data set are functions of the structural parameters of the DSGE model and are therefore restricted. An II test is provided by comparing the VAR estimated on the actual data with the VAR estimated using data simulated from the estimated DSGE model where the original data have been used to estimate the DSGE model. Le et al. (2011, 2016) propose using a Wald test based on the VAR coefficients. Alternatively, the test could be based on other features of the model, such as the associated impulse response functions or the moments - Minford, Wickens and Xu (2016) compare the test with these features and find, mostly, that the properties are quite similar.

The estimation and evaluation of misspecified DSGE models using indirect inference has been addressed previously in Dridi, Guay and Renault (2007).[2] These authors propose a two-step procedure to achieve both objectives: estimation and 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. They derive the asymptotic distribution of the test statistic under the hypothesis that the DSGE model is misspecified.

Our paper differs in several respects from that of Dridi et al. First, our focus is not on the whole DSGE model but on subsets of equations of the whole model as our aim is not to evaluate the whole model which, as Dridi et al. note, is likely to be rejected, but to find which parts of the model, if any, are not rejected. We regard this as a useful and positive exercise rather than yet another negative verdict on DSGE models. As a result we use a limited information and not a full information test. A second difference is that we derive a numerical finite sample distribution of our test statistic obtained through simulation rather than the

[1] Wickens (2014) suggested that the reason why DSGE models forecast no better than, and very similar to, a VAR could probably be attributed to the structure of the solution of a DSGE model. In order for a DSGE model to forecast better than a VAR it would be necessary to be able to forecast the future values of the exogenous variables that appear in its solution otherwise the DSGE forecasts would be, in effect, forecasts from a VAR in the endogenous variables whose coefficients are restricted by the model. And it is well-known that an unrestricted VAR is likely to forecast better than a restricted VAR, especially if the restrictions are incorrect. In effect, an unrestricted VAR is partially correcting for this.

[2] We are grateful to the referees for drawing our attention to the paper by Dridi et al.

use of asymptotic distribution theory, and hence the sandwich formulae. Third, we examine the performance of our test and compare it with that of Dridi et al. who do not evaluate the performance of their test statistic. Fourth, we use our test to carry out further empirical analysis of the widely-used Smets and Wouters (2007) DSGE model. Our main empirical finding is that some parts of the model are rejected but others are not rejected.

Our test statistic is a Wald test which Le et al. (2016) have shown is very powerful; falsifying the DSGE model's coefficients by as little as 7% usually results in a 100% rejection. In view of this whole-model finding, the non-rejection of a subset of the DSGE model would be a strongly positive result. As the solution of a DSGE model is a restricted VAR, a natural way to modify indirect inference for testing a subset of DSGE structural equations is to augment this subset with unrestricted VAR equations derived from the solution to the corresponding, but not necessarily fully specified, complete DSGE model. These unrestricted VAR equations together with the specified subset of DSGE equations form a new complete, but limited information, DSGE model. The solution to this completed DSGE model will be a VAR that incorporates the restrictions from only the specified subset of equations. Indirect inference can now be carried out as before by simulating data from the completed DSGE model and comparing the estimates of the unrestricted VAR based on the simulated and the original data.

We apply this test to two subsets of equations of the Smets and Wouters model (hereafter SWUS), namely, the wage-price sector and the expenditure equations, the consumption-investment sector. The model is based on post-war data. There is particular interest in the wage-price subset as II tests of the complete SWUS model by Le et al. (2011) reject the model, but a modified version of the model that specifies greater price flexibility is not rejected for the Great Moderation period. One of the principal aims of the SWUS model is to modify the real business cycle model by including sticky prices on the grounds that this may be why the RBC model is usually rejected by the data. The wage-price sector is therefore a critical part of the SWUS model. We find that the limited information II test confirms our original conclusion that the wage-price sector is misspecified. We also find that, although the consumption-investment sector is not rejected during the period of the Great Moderation 1980Q1-2004Q4, it is rejected for the whole sample 1947Q1-2004Q4.

After stating the relevant theory for the testing of limited information DSGE models - for the details of testing full DSGE models see Le et al (2011) - we examine the size and power properties of our limited information test of a subset of equations and we compare these with the properties of a full information test of the subset obtained by simulating all of the endogenous variables using the whole structural model. We also compare the properties of our limited information Wald test using its numerical finite sample distribution with the asymptotic distribution proposed by Dridi et al. We find that the size of our test is generally accurate (and can be made accurate by adjustment), and that its power for testing the misspecification of the wage-price sector is higher than that of Dridi et al. The power of our test is also greater for the wage-price sector than for the consumption-investment sector.

The remainder of the paper is organized as follows. Section 2 describes the limited information DSGE model consisting of a fully specified subsector of the full DSGE model together with unrestricted VARs for the other equations. Section 3 explains Indirect Inference and the testing procedure. Section 4 evaluates the performance of our limited information test for the wage-price and consumption-investment sectors of the SWUS model based on using the numerical small sample distributions of the test statistic derived from Monte Carlo simulations. In this section we also evaluate the properties of the Dridi et al. test based on asymptotic distribution theory. In Section 5 we carry out further tests of the SWUS model using both the complete sample based on post-war US data employed by SWUS and data for the Great Moderation sub-period. Section 6 concludes.

2 The limited information 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 et al. (2007) are satisfied, the solution to this model can be represented by a VAR of the 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.

Consider a subset of the structural equations of this complete DSGE model in which y_{1t} are the endogenous variables that are (partially) determined in this subset and y_{2t} are the remaining endogenous variables where $y_t = (y_{1t}, y_{2t})'$. The subset of structural equations may be written

$$A_{01}E_t y_{1t+1} + A_{02}E_t y_{2t+1} = A_{11}y_{1t} + A_{12}y_{2t} + B_1 z_t. \tag{3}$$

Where variables are not included in this subset the corresponding elements are zero.

In order to estimate or to test this subset of equations we augment the subset with unrestricted versions of the solution to y_{2t} derived from the full DSGE model for which only the subset of interest is assumed to be specified; the structural equations for y_{2t} are not specified. This gives the completed model

$$\begin{aligned} \begin{bmatrix} A_{01} & A_{02} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} E_t y_{1,t+1} \\ E_t y_{2,t+1} \\ E_t z_{t+1} \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} & B_1 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ z_t \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 \\ F_{21}^U & F_{22}^U & F_{23}^U \\ 0 & 0 & R \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ z_{t-1} \end{bmatrix} \\ &+ \begin{bmatrix} 0 & 0 & 0 \\ G_{21}^U & G_{21}^U & G_{23}^U \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \xi_{1t} \\ \xi_{2t} \\ \varepsilon_t \end{bmatrix}. \end{aligned} \tag{4}$$

where the superscript U denotes that the matrix is unrestricted.[3] This completed model is, in effect, a DSGE model with limited information. The solution to this limited information DSGE model is also a VAR but with coefficient restrictions reflecting only the structural restrictions in the subset of equations.

A special case of the DSGE model is where all of the exogenous variables

[3] In practice, these matrices are estimated by OLS from eq(2).

are unobservable and may be regarded as structural shocks. An example is the SWUS model to be examined below. This case and its solution can be represented as above both for the complete DSGE model and the limited information DSGE model.

3 Indirect inference

A full explanation of how to test a DSGE model using the method of indirect inference is given in Le et al. (2011, 2016). This can be applied to tests of a subset of the structural equations of a DSGE model by using the completed model described above which, as shown, is an equivalent representation of a limited information DSGE model. For a fully specified DSGE model, we bootstrap N samples of simulated data from the model. We then estimate the auxiliary model formed from the solution to the DSGE model which we represent as a VAR(1) using both the actual data and the N samples of simulated data. We denote the vector of all of the VAR coefficients by a and the two sets of estimates by a_T and a_S $(S = 1, ..N)$, respectively. We then use a Wald statistic (WS) based on the difference between a_T, the estimates of the VAR coefficients derived from actual data, and $\overline{a_S(\theta_0)}$, the mean of their distribution based on the simulated data. The test statistic is

$$WS = (a_T - \overline{a_S(\theta_0)})' W(\theta_0)^{-1} (a_T - \overline{a_S(\theta_0)}) \tag{5}$$

where θ_0 is the vector of parameters of the DSGE model under the null hypothesis that it is true and $W(\theta_0)$ is the weighting matrix. $W(\theta_0)$ can be obtained from the variance-covariance matrix of the distribution of simulated estimates a_S

$$W(\theta_0) = \frac{1}{N} \Sigma_{s=1}^{N} (a_s - \overline{a_s})'(a_s - \overline{a_s}) \tag{6}$$

where $\overline{a_s} = \frac{1}{N} \Sigma_{s=1}^{N} a_s$.

Asymptotically, the test statistic WS has a $\chi^2(r)$ distribution, with the number of restrictions equal to the number of elements in a_T. We carry out the test based on the empirical distribution of WS which can be obtained by bootstrapping. Appendix A shows the steps involved in finding the Wald statistic. A detailed description of the II Wald test can also be found in Le et al. (2016).

In applying this test procedure to a subset of structural equations of a DSGE model we simulate the data using the completed limited information DSGE model, equation (4) with unrestricted reduced form equations for y_2 in place of structural equations. We may then carry out the test as above; in practice as explained in Le et al. (2016) we use a VAR(1) in a limited set of y variables to obtain an appropriate level of power.

4 Evaluating the test in small samples

The size and power properties of this subset test for small samples may be examined using Monte Carlo experiments. The sample size is chosen as 200, which is typical for macro data. We design Monte Carlo simulations following the approach in Le et al. (2016). We use the SWUS model as a full DSGE model and create 1000 samples from this model, which is assumed to be true. Then we obtain from these samples the distribution of the Wald statistic by bootstrapping (the bootstrap number is 500) when the model is true. We use this distribution to assess how many times the x% False model is rejected with 95% confidence.

The false models are generated as follows. We fix the VAR coefficients of the endogenous variables that are not of interest and falsify the coefficients in the structural equations in the subset that is to be tested. We generate the falseness by introducing an increasing degree of mis-specification for the parameters of the subset of interest. We introduce two types of falseness: altering the parameters in the subset by +/- x% alternating (even-numbered parameters positive, odd ones negative) and +/- x% randomly. The level of falseness (x%) is increased from 1% to 3%, 5%, 7%, 10%,15% and 20%. In this way we construct a False DSGE model whose parameters are moved x% away from their true values in both directions.[4]

We are interested in (a) the accuracy of inferences based on the test when the subset model is false, and (b) whether the subset model is causing the whole model to be rejected. In (a) we focus only on the parameters of the subset using the unrestricted VAR solutions for the rest of the model. We refer to this as the limited information test of a subset of equations. In (b) our test of the subset model takes account of the parameter restrictions in the whole model. We refer to this as a full information test of a subset of equations.

[4] See Le et al. (2016) section 4.1 for full details of the experiments. For all the experiments, the eigenvalues of reduced-form VAR coefficients are all strictly less than unity in modulus, so the Fernandez-Villaverde et al. (2007) condition that the DSGE model has a VAR representation is satisfied.

4.1 Test of the wage-price equations

We begin with the wage-price equation subset of the SWUS model. These equations are derived under the assumption of Calvo contracts as New-Keynesian Phillips curves, respectively for wage and price-setting as:

$$\begin{aligned} w_t &= w_1 w_{t-1} + (1-w_1)(E_t w_{t+1} + E_t \pi_{t+1}) - w_1 \pi_t + w_3 \pi_{t-1} \\ &\quad - w_4 (w_t - (\sigma_l l_t + \frac{1}{1-\lambda/\gamma}(c_t - \lambda/\gamma c_{t-1})) + \varepsilon_t^w. \end{aligned} \tag{7}$$

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 E_t \pi_{t+1} - \pi_3 (\alpha(k_t^s - l_t) - w_t) + \varepsilon_t^p. \tag{8}$$

The wage and price mark up disturbances follow an AR(1) process: $\varepsilon_t^w = \rho_w \varepsilon_{t-1}^w + \eta_t^w, \varepsilon_t^p = \rho_p \varepsilon_{t-1}^p + \eta_t^p$.[5]

The two key endogenous variables are w_t and π_t. The other endogenous variables in the full SWUS model are $y_t, c_t, i_t, l_t, w_t, k_t^s, q_t$. There are two exogenous shocks ε_t^w and ε_t^p. The equations for the two variables (w_t and π_t) form our 'subset model'.

Table 1: Rejection rates of the wage-price sector at 5% significant level.

Falseness is given by +/- x% alternation								
	0%	1%	3%	5%	7%	10%	15%	20%
LIST unadjusted	0.053	0.077	0.355	0.928	1.000	1.000	1.000	1.000
LIST adjusted	0.050	0.076	0.345	0.923	1.000	1.000	1.000	1.000
FIST	0.050	0.074	0.385	0.973	1.000	1.000	1.000	1.000
Falseness is given by +/- x% randomly								
	0%	1%	3%	5%	7%	10%	15%	20%
LIST unadjusted	0.056	0.117	0.176	0.574	0.727	0.801	1.000	1.000
LIST adjusted	0.050	0.070	0.126	0.454	0.630	0.731	0.999	1.000
FIST	0.050	0.074	0.136	0.402	0.568	0.798	0.999	1.000

Notes: These are rejection rates under different types and degrees of parameter falsification; LIST unadjusted and LIST adjusted denote limited information subset tests without and with size adjustment; FIST denotes full information subset tests. Auxiliary VAR(1) has 2 variables w_t and π_t.

Table 1 gives the size and power of the tests for the wage-price subset of equations using an auxiliary model that is a VAR(1) in the two variables explained in the subset, w_t and π_t. The first two rows report the power of the

[5] See Appendix B for the structural parameters values that are used in simulation. See Smets and Wouters (2007) for the full model and details of the two equations, including an explanation of the parameters.

limited information subset test (LIST) where the equations that are not part of the subset are replaced in the simulations by an unrestricted VAR. The first row gives the rejection rates with the different degrees of falseness. The entry in this row shows that the sizes of the test are 5.3% and 5.6% and not 5%. As this makes the test a little too conservative, and so slightly distorts the power of the test, in the second row we adjust for this by scaling the rejection rates to give a size of test of 5%. In the third row we report the power of the subset test where all of the variables are simulated using the structural restrictions in the whole model. In the power calculations for the subset test the structural coefficients in the equations that are not part of the subset are held constant in all simulations. Used as a test - instead of, as here, to examine the potential loss of power due to using unrestricted VARs to simulate data for the endogenous variables in the other structural equations - this would generate a test of the whole model. We refer to these power calculations as full information subset tests (FIST). For both the LIST test and FIST power calculations only the restrictions in the subset of equations are tested.

As expected, the power of the LIST test is slightly reduced after adjustment for its size. The power itself is very high. For the first type of falsification rejection is virtually guaranteed for degrees of falsity above 3%; the random falsification model will be rejected more than 70% of the time for 10% falsification. The power of the LIST test and the FIST power calculations are remarkably close. This is an important result as it shows that the loss of information in the limited information test has almost no effect on the power of the test.

We now repeat the analysis using the same auxiliary model that we used to test the whole model: a three variable VAR(1) in the key variables y_t (output), π_t (inflation), r_t (interest rate). In Table 2 we report results for the adjusted LIST test and the FIST power calculations based on one type of falsification.

Table 2: Rejection rates of the wage-price sector at 5% significant level.

Falseness is given by +/- x% alternation								
	0%	1%	3%	5%	7%	10%	15%	20%
LIST adjusted	0.050	0.058	0.149	0.328	0.644	1.000	1.000	1.000
FIST	0.050	0.058	0.158	0.455	0.775	0.851	0.930	0.978

Notes: These are rejection rates under different degrees of parameter falsification; LIST adjusted denotes limited information subset tests with size adjustment; FIST denotes full information subset tests. Auxiliary VAR(1) has 3 variables y_t, π_t and r_t.

The power remains high and there is little difference in the powers of the limited and full information tests from the addition to the auxiliary VAR of variables that are not in the subset being tested.

4.2 Test of the consumption-investment equations

The dynamics of consumption are derived from the consumption Euler equation and those of investment from the investment Euler equation. The equations are, respectively,

$$c_t = c_1 c_{t-1} - (1-c_1)E_t c_{t+1} + c_2(l_t - E_t l_{t+1}) - c_3(r_t - E\pi_{t+1} + \varepsilon_t^b). \tag{9}$$

$$i_t = i_1 i_{t-1} + (1-i_1)E_t i_{t+1} + i_2 q_t + \varepsilon_t^i. \tag{10}$$

The consumption and investment mark-up disturbance follows an AR(1) process: $\varepsilon_t^b = \rho_b \varepsilon_{t-1}^b + \eta_t^b$, $\varepsilon_t^i = \rho_i \varepsilon_{t-1}^i + \eta_t^i$. The two key endogenous variables are c_t and i_t. The other endogenous variables in the full SWUS model, namely $y_t, l_t, w_t, r_t, \pi_t, k_t^s$, q_t, are generated from the unrestricted VAR.

Table 3 replicates Table 1 for the consumption investment sector. We begin our analysis by testing the subset consisting of only the two variables (c_t and i_t), using a VAR(1) in these two variables as the auxiliary model.

Table 3: Rejection rates of the consumption-investment sector at 5% level.

Falseness is given by +/- x% alternation								
	0%	1%	3%	5%	7%	10%	15%	20%
LIST unadjusted	0.095	0.102	0.171	0.202	0.224	0.286	0.425	0.424
LIST adjusted	0.050	0.051	0.086	0.094	0.123	0.174	0.311	0.288
FIST	0.050	0.055	0.057	0.056	0.056	0.069	0.103	0.140
Falseness is given by +/- x% randomly								
	0%	1%	3%	5%	7%	10%	15%	20%
LIST unadjusted	0.096	0.160	0.231	0.066	0.057	0.378	0.332	0.253
LIST adjusted	0.050	0.086	0.138	0.035	0.032	0.286	0.215	0.175
FIST	0.050	0.060	0.059	0.056	0.132	0.181	0.124	0.308

Notes: These are rejection rates under different types and degrees of parameter falsification; LIST unadjusted and LIST adjusted denote limited information subset tests without and with size adjustment; FIST denotes full information subset tests. Auxiliary VAR(1) has 2 variables c_t and i_t.

The unadjusted size of the limited information test is 9.5%. This is less accurate than for the wage-price sector and a little more conservative. The adjusted limited information test statistic for the consumption-investment sector is much lower than that for the wage-price sector. This suggests that the consumption-investment sector plays a considerably smaller role in the rejection of the whole SWUS model than the wage-price sector. Once again there is not a large difference in the powers of the limited and full information tests.

In Table 4 we repeat Table 2 for the consumption-investment sector using the same auxiliary model as we used to test the whole model: a three variable VAR(1) in the key variables y_t (output), π_t (inflation), r_t (interest rate).

Table 4: Rejection rates of the consumption-investment sector at 5% level.

Falseness is given by +/- x% alternation								
	0%	1%	3%	5%	7%	10%	15%	20%
LIST adjusted	0.050	0.054	0.058	0.050	0.054	0.056	0.067	0.068
FIST	0.050	0.055	0.061	0.088	0.111	0.171	0.337	0.511

Notes: These are rejection rates under different degrees of parameter falsification; LIST adjusted denotes the limited information subset test with size adjustment; FIST denotes full information subset tests. Auxiliary VAR(1) has 3 variables y_t, π_t and r_t.

The power of the LIST test is still very weak, as is the power of the FIST test, which uses the whole model: falsifying the parameters of this subset generates low rejection rates that hardly increase with rising falsity. Again this indicates that the consumption-investment subsector of the model will not cause the whole model to be rejected unless it is extremely false (15% or more).

4.3 Robustness of the test

We consider a number of issues concerning our procedures and their effects on the properties of the test. First, we discuss the effect of using an alternative falseness criterion of the structural parameters when evaluating power. Second, we compare our previous tests which were based on a numerical finite sample distribution of the test statistic with a test that uses the asymptotic distribution of the Wald statistic obtained by Dridi et al. (2007).

Falsifying the parameters by +/- x% of standard deviation

In order to reflect that some structural parameters are more precisely estimated than others, it is of interest to falsify the parameters in the power calculation in terms of percentages of the standard deviations of the estimates of the structural parameters. We therefore repeat the earlier analysis using the same auxiliary model as we used to test the whole model: a three variable VAR(1) in the key variables y_t (output), π_t (inflation), r_t (interest rate), by falsifying the parameters by +/- x% of their standard deviation. The results are reported in Table 5.

Table 5: Rejection rates of the wage-price and cons-invest sectors at 5% level.

Wage-price sector								
	0%	3%	7%	10%	15%	20%	25%	35%
LIST adjusted	0.05	0.059	0.104	0.186	0.396	0.675	0.900	1.000
FIST	0.05	0.042	0.090	0.168	0.456	0.802	0.954	1.000
Consumption-investment sector								
	0%	3%	7%	10%	15%	20%	25%	35%
LIST adjusted	0.05	0.044	0.039	0.035	0.034	0.033	0.030	0.030
FIST	0.05	0.052	0.061	0.064	0.071	0.077	0.098	0.144

Notes: These are rejection rates under different degrees of falsification; falsification is by +/- x% of the standard deviation; LIST adjusted denotes the limited information subset test adjusted for size; FIST denotes full information subset test. Auxiliary VAR(1) has 3 variables y_t, π_t and r_t.

The results are consistent with the previous simulations based on falsifying by +/- x% from the true parameters. The power of the test is high for the wage-price subset, but is low for the consumption-investment subset. Using this alternative falsification method does not therefore seem to affect the results greatly.

Using the asymptotic distribution

In our numerical finite sample results we used a bootstrapped weighting matrix. This originated in classical minimum distance estimation theory; see, for example, Hall et al. (2012), Le et al. (2011) and Guerron-Quintana, Inoue and Kilian (2017). The latter showed that II estimation based on bootstrapping the weighting matrix gives a consistent estimator, but its asymptotic distribution is not a standard Normal. This is why using the numerical finite sample distribution is advisable. In previous papers, such as Meenagh et al. (2016), we have also found that it gives a test with good power.

Nonetheless, it is of interest to compare the use of the asymptotic distribution of the Wald statistic based on the sandwich formula for the weighting matrix with the numerical finite sample distribution. Dridi, Guay and Renault (2007) show that the weighting matrix based on asymptotic distribution theory is[6],

$$\begin{aligned} W(\theta_0) &= J_0^{-1} I_0 J_0^{-1} + \frac{1}{S} J_0^{*-1} I_0^{*} J_0^{*-1} + \left(1 - \frac{1}{N}\right) J_0^{*-1} K_0^{*} J_0^{*-1} \\ &\quad - J_0^{-1} K_0 J_0^{*-1} - J_0^{*-1} K_0' J_0^{*-1}. \end{aligned} \tag{11}$$

We repeat the analysis using the same auxiliary model that we used to test the whole model: a three variable VAR(1) in the key variables. We compare the small sample performance of the adjusted LIST test which is computed using bootstrapping with the asymptotic weighting matrix that employs the sandwich formula. The results for the wage-price and consumption-investment sectors are reported in Table 6. The LIST adjusted results are taken from Tables 2 and 4.

Table 6: Rejection rates of the wage-price and cons-invest sectors at 5% level.

Wage-price sector								
Weighting matrix	0%	1%	3%	5%	7%	10%	15%	20%
LIST adjusted	0.05	0.058	0.149	0.328	0.644	1.000	1.000	1.000
Asymptotic	0.05	0.060	0.071	0.123	0.834	1.000	1.000	1.000
Consumption-investment sector								
Weighting matrix	0%	1%	3%	5%	7%	10%	15%	20%
LIST adjusted	0.05	0.054	0.058	0.050	0.054	0.056	0.067	0.068
Asymptotic	0.05	0.054	0.055	0.056	0.056	0.057	0.062	0.060

Notes: These are rejection rates under different degrees of falsification; falsification is by +/- x% alternation; LIST adjusted denotes the limited information subset test adjusted for size; Asymptotic denotes the limited information subset test using the asymptotic weighting matrix. Auxiliary VAR(1) has 3 variables y_t, π_t and r_t.

We find that for the wage-price sector the power of the numerical finite sample distribution is higher for low levels of falseness, but the two are similar for higher levels of falseness. This suggests that the asymptotic approximation improves as the degree of falsity increases but is nowhere superior to the finite sample distribution. For the consumption-investment sector there is little difference in the results. This suggests that the structural coefficients in this subset are less well determined.

5 An application to US data

One of the purposes of this test is to throw more light on individual sectors of a DSGE model: which sectors are not rejected and how they might be respecified

[6] See Appendix C for details of how to derive the sandwich form weighting matrix $W(\theta_0)$.

if they are rejected. This is more constructive than a blanket rejection of the whole model.

In their work evaluating the Smets and Wouters model on US post-war data, Le et al. (2011) found that the model as estimated by SWUS with Bayesian methods was decisively rejected by the II test on the full post-war sample. Le et al also found that a ‘New Classical’ version of the model was also decisively rejected. They then considered a compromise version in which a competitive sector coexisted with a sticky-price sector in both the labour and goods markets. They found that, when re-estimated by II, although the hybrid model was still rejected using the whole sample, it was not rejected for the Great Moderation period. This suggested that the original version of the wage-price subset may be a problematic component of the SWUS model. It is therefore of interest to revisit this issue using the new subset test proposed in this paper by comparing the performance of wage-price and consumption-investment sectors for the whole post-war sample used by Smets and Wouters from 1947Q1-2004Q4 and for the period of the Great Moderation 1980Q1-2004Q4.

In Table 7 we report the p-values for the adjusted LIST test and for the asymptotic LIST test for the two sectors and sample periods.

Table 7: LIST tests of the Smets-Wouters model for the two sample periods.

	Great Moderation 1980Q1-2004Q4		Post War 1947Q1-2004Q4	
Test	Wage-price	Cons-invest	Wage-price	Cons-invest
LIST adjusted	0.000	0.140	0.000	0.003
Asymptotic	0.143	0.218	0.003	0.021

Notes: These are p-values for limited information subset tests of the joint wage-price and consumption-investment sectors. LIST adjusted denotes the limited information subset test adjusted for size; Asymptotic denotes the limited information subset test using the asymptotic weighting matrix. Auxiliary VAR(1) has 3 variables y_t, π_t and r_t.

Using the limited information test we reject the wage-price sector for both samples and we reject the consumption-investment sector for the whole sample. These results support our conjecture that during the period of the Great Moderation the problem did indeed lie with the wage-price sector and not the spending sector. It also suggests that there may have been structural change in the consumption-investment sector before and after the Great Moderation.

The asymptotic test seems to have lower power, especially in the Great Moderation period as, in that sub-period, neither sector is rejected. This result is consistent with the lower power of the asymptotic test reported above. A contributory factor may be the smaller sample size of the Great Moderation period.

6 Conclusion

Testing DSGE models is not common, especially if they are estimated by Bayesian methods. Previously tests of DSGE models have been of the whole model and they have commonly rejected the model. A more positive approach would be to examine whether some parts of the DSGE model are not rejected. This would help to identify which parts of the full model need to be respecified. This paper suggests a way of carrying out tests on a subset of equations of a DSGE model that is based on applying limited information theory via indirect inference. The test is a modification of the II test proposed by Le et al. (2011, 2016) for complete DSGE models and draws on the theory of estimating and testing limited information models.

The main drawback of any limited information test is that by not using the structural information in the whole model the power of the test of a subset of the structural coefficients may be lower. This is only likely, however, if the rest of the model is correctly specified. If it is not then such misspecification may contaminate the estimates of the structural coefficients in the subset under test and hence affect its power. A limited information test may therefore be more robust against misspecification elsewhere in the model.

Our limited information test has the added advantages of being easy to perform and seems to have good power in small samples; better, in fact, than the more commonly used asymptotic distribution theory, and similar to a test based on using the structural information in the whole model instead of just the subset of equations under test. The properties of both our limited information test and an asymptotic test are examined for the wage-price and consumption-investment subsectors. We also found that calculating the power of the tests using parameter deviations based on the standard deviations of the structural parameters rather than the parameters themselves had little effect on our initial results. These results are robust to the variables included in the auxiliary model; this seems to reflect the fact that all are variant approximations of the same reduced form.

The application of these tests to the wage-price and consumption-investment sectors of the SWUS model is based on the finding in Le et al. (2011) that the wage-price sector is misspecified, having excessive stickiness. The particular interest in this finding is that one of the principal aims of the SWUS model was to modify the real business cycle model by including sticky prices on the grounds that this may be why the RBC model was usually rejected by the data. We found that, using the whole sample of US post-war data, and for the sub-period of the Great Moderation, the whole model is rejected. Our new result obtained by using our limited information test is that for the Great Moderation, while the wage-price sector is rejected, the consumption-investment sector is not rejected. This suggests that the rejection of the whole SWUS model over the Great Moderation period may be due to the wage-price sector in the model. We conclude that this subset test may prove to be both a useful tool for analysing empirical weaknesses in DSGE models and a constructive tool for isolating which parts of the model need to be respecified.

References

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

[2] Fernandez-Villaverde, J., Rubio-Ramirez, F., Sargent, T, and Watson, M. (2007). 'ABCs (and Ds) of Understanding VARs', American Economic Review, 1021-1026.

[3] Godfrey, L., and Wickens, M. (1982). 'A simple derivation of the limited information maximum likelihood estimator', Economics Letters,10(3-4,) 277-283.

[4] Gourieroux, C., A. Monfort and E. Renault. (1993). 'Indirect inference', Journal of Applied Econometrics, Volume 8: 85-118.

[5] Gregory, A. and Smith, G. (1993). Calibration in macroeconomics, in: Maddala, G. (Ed.), Handbook of Statistics vol. 11, Elsevier, St. Louis, Mo., pp. 703–719.

[6] Guerron-Quintana, P., Inoue, A. and Kilian, L. (2017). 'Impulse response matching estimators for DSGE models', Journal of Econometrics, 196(1): 144-155.

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

[8] 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.

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

[10] McCallum, B.T.(1976). 'Rational expectations and the natural rate hypothesis: some consistent estimates', Econometrica, 44, 4–52.

[11] Meenagh, D, Minford, P, Wickens, M and Xu, Y. (2016). 'What is the truth about DSGE models? Testing by indirect inference', Cardiff economic working paper, E2016/14, http://carbsecon.com/wp/E2016_14.pdf

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

[13] Phillips, P.C.B. and Wickens, M. (1978). Exercises in Econometrics, Volume II, Philip Allan/Ballinger.

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

[15] Smith, A. A. (1993). 'Estimating nonlinear time-series models using simulated vector autoregressions', Journal of Applied Econometrics, 8, S63-84.

[16] Wickens, M.R. (1982). 'The efficient estimation of econometric models with rational expectations', Review of Economic Studies, 49, 55–67.

[17] Wickens, M. (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 SW 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

$$W(\theta_0) = \frac{1}{N}\Sigma_{k=1}^{N}(a_k - \overline{a_k})'(a_k - \overline{a_k}) \tag{12}$$

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

$$WS = (a_T - \bar{a}_s(\theta_o))'W(a_s(\theta_o))^{-1}(a_T - \bar{a}_s(\theta_o)) \tag{13}$$

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 subsets of equations

The parameters in the wage-price and consumption-investment subsets of the Smets-Wouters (2007) model are:

$\pi_1 = \iota_p/(1+\beta\gamma^{1-\sigma_c}\iota_p)$;
$\pi_2 = \beta\gamma^{1-\sigma_c}/(1+\beta\gamma^{1-\sigma_c}\iota_p)$;
$\pi_3 = 1/(1+\beta\gamma^{1-\sigma_c}\iota_p)[(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/(\xi_p((\phi_p-1)\varepsilon_p+1))]$.
$w_1 = 1/(1+\beta\gamma^{1-\sigma_c})$;
$w_2 = (1+\beta\gamma^{1-\sigma_c}\iota_w)(1+\beta\gamma^{1-\sigma_c})$;
$w_3 = \iota_w/(1+\beta\gamma^{1-\sigma_c})$;
$w_4 = 1/(1+\beta\gamma^{1-\sigma_c})[(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/(\xi_w((\phi_w-1)\varepsilon_w+1))]$.
$c_1 = (\lambda/\gamma)\,(1+\lambda/\gamma)$;
$c_2 = [(\sigma_c-1)(W_*^h L_*/C_*)\ /[(1+\lambda/\gamma)\sigma_c]$;
$c_3 = (1-\lambda/\gamma)\,[(1+\lambda/\gamma)\sigma_c]$.
$i_1 = 1/(1+\beta\gamma^{1-\sigma_c})$;
$i_2 = 1/(1+\beta\gamma^{1-\sigma_c})\gamma^2\varphi$.

In the power calculations the true values assumed for these parameters are based on the posterior modes: see Smets and Wouters (2007) Tables 1A and1B. They are listed below. The false models are generated by moving the free parameters in the relevant subset +/-x% away from their true values.

Fixed parameters		
β	0.9975	Discount factor
δ	0.025	Depreciation rate of capital
ε_p	10	Curvature of Kimball goods aggregator
ε_w	10	Curvature of Kimball labour aggregator
ϕ_w	1.5	1+ steady state wage mark up
ϕ_p	1.6	1+ steady state price mark up

Note: Quarterly data are used.

Free parameters

	Posterior Mode	St. Dev.		Posterior Mode	St. Dev.
φ	5.48	1.50	σ_a	0.45	2.00
σ_c	1.39	0.37	σ_b	0.24	2.00
h	0.71	0.10	σ_g	0.52	2.00
ξ_w	0.73	0.10	σ_I	0.45	2.00
σ_l	1.92	0.75	σ_r	0.24	2.00
ξ_p	0.65	0.10	σ_p	0.14	2.00
ι_w	0.59	0.15	σ_w	0.24	2.00
ι_p	0.22	0.15	ρ_a	0.95	0.20
ψ	0.54	0.15	ρ_b	0.18	0.20
Φ	1.61	0.12	ρ_g	0.97	0.20
r_π	2.03	0.25	ρ_I	0.71	0.20
ρ	0.81	0.10	ρ_r	0.12	0.20
r_y	0.08	0.05	ρ_p	0.90	0.20
$r_{\triangle y}$	0.22	0.05	ρ_w	0.97	0.20
π	0.81	0.10	μ_p	0.74	0.20
l	-0.10	2.00	μ_w	0.88	0.20
$\bar{\gamma}$	0.43	0.10			
α	0.19	0.05			

Note: St. Dev. denotes the standard deviation; $\bar{\gamma} = 100(\gamma - 1)$ is the common quarterly trend growth rate to real GDP.

Appendix C: 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:

$$\begin{aligned}
I_0^* &= \frac{1}{N}\sum_{s=1}^{N} E(S_t^s(Y^{0s},X^s)S_t^s(Y^{0s},X^s)') \\
J_0^* &= -\frac{1}{N}\sum_{s=1}^{N} E(H_t^s(Y^{0s},X^s)) \\
K_0 &= \frac{1}{N}\sum_{s=1}^{N} E(S_t(Y^0,X)S_t^*(Y^{0*},X^*)) \\
K_0^* &= \frac{1}{N(N-1)/2}\sum_{s\neq l}^{N} E(S_t^s(Y^{0s},X^s)S_t^l(Y^{0l},X^l))
\end{aligned} \tag{14}$$

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

$$\sqrt{T}(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0) \backsim N(\mathbf{0},\Xi(N,W) \tag{15}$$

with

$$\Xi(N,W)=\left\{\frac{\partial'(\boldsymbol{\delta})}{\partial(\theta_0)}W(\theta_0)^{-1}\frac{\partial'(\boldsymbol{\delta})}{\partial(\theta_0)}\right\}^{-1}. \tag{16}$$

and an asymptotic weighting matrix

$$\begin{aligned}
W(\theta_0) &= J_0^{-1}I_0J_0^{-1}+\frac{1}{N}J_0^{*-1}I_0^*J_0^{*-1}+\left(1-\frac{1}{N}\right)J_0^{*-1}K_0^*J_0^{*-1} \\
&\quad -J_0^{-1}K_0J_0^{*-1}-J_0^{*-1}K_0'J_0^{*-1}
\end{aligned} \tag{17}$$

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

The II Wald statistic for a DSGE model is given by

$$WS = (\hat{\boldsymbol{\delta}}_T - \bar{\boldsymbol{\delta}}_s(\theta_0))' W(\theta_0)^{-1} (\hat{\boldsymbol{\delta}}_T - \bar{\boldsymbol{\delta}}_s(\theta_0)) \tag{18}$$

where $\hat{\boldsymbol{\delta}}_T$ is the ML estimation on the coefficients of the VAR using the actual data and $\bar{\boldsymbol{\delta}}_s(\theta_0)$ is the mean of the ML estimation on the coefficients of VAR using simulated data. $W(\theta_0)$ is the weighting matrix, which can obtained either from bootstrap samples in equation (6) or from the asymptotic variance in equation (17).