Caglar, Evren; Chadha, Jagjit S.; Shibayama, Katsuyuki

Working Paper

Bayesian Estimation of DSGE Models: Is the Workhorse Model Identified?

Working Paper, No. 1205

Provided in Cooperation with:
Koç University - TÜSİAD Economic Research Forum, Istanbul

Suggested Citation: Caglar, Evren; Chadha, Jagjit S.; Shibayama, Katsuyuki (2012) : Bayesian Estimation of DSGE Models: Is the Workhorse Model Identified?, Working Paper, No. 1205, Koç University-TÜSİAD Economic Research Forum (ERF), Istanbul

This Version is available at:
https://hdl.handle.net/10419/108589

KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM
WORKING PAPER SERIES

BAYESIAN ESTIMATION OF DSGE MODELS: IS THE WORKHORSE MODEL IDENTIFIED?

Evren Caglar
Jagjit S. Chadha
Katsuyuki Shibayama

Working Paper 1205
February 2012

KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM
Rumelifeneri Yolu 34450 Sarıyer/Istanbul

Bayesian Estimation of DSGE models: Is the Workhorse Model Identified?*

Evren Caglar† Jagjit S. Chadha‡and Katsuyuki Shibayama§

University of Kent

November 2011

Abstract

Koop, Pesaran and Smith (2011) suggest a simple diagnostic indicator for the Bayesian estimation of the parameters of a DSGE model. They show that, if a parameter is well identified, the precision of the posterior should improve as the (artificial) data size T increases, and the indicator checks the speed at which precision improves. It does not require any additional programming; a researcher just needs to generate artificial data and estimate the model with different T. Applying this to Smets and Wouters' (2007) medium size US model, we find that while exogenous shock processes are well identified, most of the parameters in the structural equations are not.

KEYWORDS: Bayesian Estimation, Dynamic stochastic general equilibrium models, Identification.

JEL CLASSIFICATION: C51, C52, E32.

*__Acknowledgements__: We thank Gary Koop, Hashem Pesaran and Ron Smith for advice and support. We also thank Keisuke Otsu for detailed comments.

†PhD candidate, School of Economics, University of Kent, Canterbury. E-mail: ec265@kent.ac.uk.

‡Chair in Banking and Finance, School of Economics, University of Kent, Canterbury and Centre for International Macroeconomics and Finance, University of Cambridge. E-mail: jsc@kent.ac.uk.

§Address for correspondence: Department of Economics, University of Kent, Canterbury, Kent, CT2 7NP, U.K. Phone: +44(0)122782-4714. E-mail: k.shibayama@kent.ac.uk

1 Introduction

Many macroeconomists have expressed concern about the extent to which identification of DSGE models may or may not have been achieved during estimation. Reflecting the rapid progress of Bayesian estimation techniques, it is now increasingly more common to estimate DSGE models than to simply calibrate them. The problem is, however, that if a parameter is not identified, this means that the data (and the prior) cannot pin down the value of this parameter, and if a parameter is only weakly identified, this means that a small change in, say, the sample variation causes a large change in the parameter estimate. Compared with standard linear identification problems in econometrics, the estimation of DSGE models involves nonlinear estimation under many theoretical parameter restrictions and so the identification is much harder.

Much worse, in the Bayesian framework the prior often masks the problem of non- or weak identification by the data.[1] That is, even if data provide little or no information of a parameter, it still can be seemingly identified solely because of its prior. Koop *et al.* (2011) discuss, from a pure Bayesian perspective, this observation may not necessarily be problematic and we might simply want to thank our informative priors. However, some (or perhaps most) researchers may regard this as rather embarrassing, as econometric-based inference may only then rely only on researchers' initial beliefs and not on the data. In this respect, Canova and Sala (2009) among others, warn against the current practice of comparing the prior and posterior densities of a parameter to check the informativeness of data: since a parameter may be identified only jointly with others and not individually, even if these densities have different shapes, still there is a significant possibility that any given parameter may be unidentified.

As a result of these problems, two strands of diagnostic indicators have been developed. The first line of indicators sets an intermediate target and investigates the Jacobian of such a target with respect to the deep parameters of a model. This line of indicators has been pioneered by Iskrev (2010a), Iskrev and Ratto (2010) and Komunjer and Ng (2009). Typically, this intermediate target is a set of data moments. If the

[1] See Canova and Sala (2009) and Koop *et al.* (2011) among others.

Jacobian of the data moments is column rank deficient, there are two possibilities; (i) one or more parameters do not affect any data moments at all; and (ii) a change in one parameter is totally offset by changes in other parameters and hence again may not affect any moments. The latter case, which is presumably more common than the former, is often referred to as partially identified or perfect collinearity among parameters. Iskrev (2010a) also proposes a check of the Jacobian of the reduced form parameters with respect to the deep parameters, so-called Iskrev's J_2.[2] Though this proposal only checks whether a necessary condition of non-identification is satisfied, as it does not rely on any data, it is often convenient especially when we have no *a priori* information about the identification. Also, it is useful to detect the source of non- or weak identification: if the Jacobian is column full rank but a parameter is only weakly identified or not at all identified, it is evident that non or weak identification is because of data limitations and not because of the model structure.

The second line of indicators, such as Koop *et al.* (2011, KPS henceforth) and Iskrev (2010b), exploits the Information matrix, which is the expectation of the Hessian. This idea is very straightforward: if the likelihood function is flat along a particular direction at a likelihood *mode,* i.e. the Hessian is singular, the value of the likelihood (or posterior density) does not change along this direction and hence there are infinitely many combinations of parameters that achieve the *maximum* likelihood. The main difference between KPS and Iskrev (2010b) is that the former is mainly interested in the identification by data, whereas Iskrev (2010b) checks the identification by both the prior and data. This point is very important and we will discuss this more deeply in our main analysis. One practical weakness of this second approach is that, as opposed to the Jacobian based methods, if the Hessian is singular, it may be hard, if not impossible, to pin down the maximum point. This is because practically most maximizing algorithm require a non-singular (i.e., strictly negative definite) Hessian; otherwise, the likelihood mode is not well defined. This Catch-22 problem seems to be common for most Hessian based approaches. Importantly this means that this class of indicators work only for

[2]In this case, the intermediate target is the coefficients of the reduced form model solved by, say, Sims' (2002) QZ method.

weakly identified parameters; a researcher has to obtain *a priori* information about the parameters that are totally unidentified before implementing this class of indicator. However, as opposed to the Jacobian-based approach, the Hessian-based approach is a *full information approach*, in the sense that it exploits the likelihood (or posterior density), which contains all the information that is available.[3]

The purpose of this article is to investigate the KPS indicator. KPS suggest two separate methods for checking the presence and strength of identification of the parameters of DSGE models. Their first indicator is based on Bayesian theory. Suppose, for example, that it is not known if a parameter is identified or not. If it is unidentified, "the marginal posterior of this parameter will equal the posterior expectation of the prior of this parameter conditional on the identified parameters". The second method, relying on the asymptotic theory, says that the precision of a parameter estimate will increase at the rate of the data size T, if it is identified. One merit of this second method lies in the simplicity of its implementation: in practice, it does not require any additional (time consuming) programming or simulations because it just examines the Hessian (or posterior variances) for (artificial) data sets with different sizes. What a researcher then has to do, when estimating any model, is simply to check the speed at which the parameter precision increases.

As the second method is more widely applicable, we apply this method to analyse the identification of a popular DSGE model of Smets and Wouters (2007) for the US. As many researchers use Smets and Wouters (2007), or its variants, as a testing ground for their identification methods, we are thus able to compare our results with theirs. Although we need to investigate other key models as well to be conclusive, broadly speaking our finding on the Smets and Wouters model is consistent with others, such as Iskrev (2010a) and Iskrev and Ratto (2010), in that we should be cautious about whether estimated parameters are indeed identified or not. In this paper, we discuss several practical issues in computing and interpreting the KPS indicator.

The rest of the paper is organized as follows: Section 2 briefly introduces the idea of

[3]Note though that both the Jacobian- and Hessian-based approaches are local rather than global indicators.

KPS and the design of our experiment, Section 3 summarizes our main findings, Section 4 is reserved for a brief discussion of the methodology of the KPS in light of our results and finally Section 5 concludes.

2 Identification based on Asymptotic Precision

2.1 The KPS Idea

For completeness, we start with reviewing the intuition of the KPS indicator.[4] Consider the Bayesian estimation of a DSGE model. Let $\theta = (\theta_1, \theta_2 ... \theta_n)$ be a parameter vector and T be the size of the data. Suppose that the posterior density is well approximated by a normal distribution. In this case, the posterior mode, $\bar{\theta}_T$, is the average of the prior mode $\underline{\theta}$ and the data likelihood mode, $\hat{\theta}_T$, weighted by their respective precision $\underline{H}$ and $T\hat{S}_T$. That is,

$$\bar{\theta}_T = \bar{H}_T^{-1}\left(T\hat{S}_T\hat{\theta} + \underline{H\theta}\right), \tag{1}$$

$$\bar{H}_T = T\hat{S}_T + \underline{H}, \tag{2}$$

where $\bar{H}_T$ is the posterior Hessian. Note that as $T \to \infty$, $\bar{H}_T^{-1}$ asymptotes to the true variance-covariance matrix of parameter estimates.

Now suppose that all parameters are identified. In this case, $T^{-1}\bar{H}_T$ converges to $\hat{S}_T$ as $T \to \infty$,

$$T^{-1}\bar{H}_T = \hat{S}_T + T^{-1}\underline{H} \to \hat{S}_T.$$

At the limit, $\hat{S}_T$ (data precision, $T\hat{S}_T$, divided by T) converges to a certain point, while the prior precision, $\underline{H}$, is dwarfed. That is, the data dominates the prior as T increases. Since $T^{-1}\bar{H}_T$ converges to a certain value, it is clear that posterior precision $\bar{H}_T$ improves at rate T.

Let us focus on one specific parameter, say, the first parameter θ_1. Under the normality assumption, its mean is $\bar{\theta}_{1T}$ and its precision, $\bar{h}_{11}$, is given as $\bar{h}_{11} =$

[4] See Koop *et al.* (2011) for a comprehensive analysis.

$\bar{H}_{11} - \bar{H}_{12}\bar{H}_{22}^{-1}\bar{H}_{21}$.[5] Hence, we get

$$T^{-1}\bar{h}_{11} = \left(\hat{S}_{11} + T^{-1}\underline{H}_{11}\right) - \left(\hat{S}_{12} + T^{-1}\underline{H}_{12}\right)\left(\hat{S}_{22} + T^{-1}\underline{H}_{22}\right)^{-1}\left(\hat{S}_{21} + T^{-1}\underline{H}_{21}\right). \quad (3)$$

Following the same analysis as above, at the limit,

$$\lim_{T\to\infty} T^{-1}\bar{h}_{11} = \hat{S}_{11} - \hat{S}_{12}\hat{S}_{22}^{-1}\hat{S}_{21} = \left(\hat{S}_{11}^{-1}\right)^{-1},$$

which is the inverse of the $(1,1)$ element of $\hat{S}_{T=\infty}^{-1}$. Since the prior is dominated at the limit, let us focus on $T\hat{S}_T$. From the standard, or frequentist, econometric theory, it is easy to see, if $\bar{\theta}_1$ is well identified, $\hat{S}_{11}^{-1}$ approaches a certain number as $T \to \infty$; in other words, the variance $T^{-1}\hat{S}_{11}^{-1}$ of $\bar{\theta}_1$ shrinks at rate T. Intuitively, this means that, when there is more data, the estimation becomes more precise. These observations lead KPS to recommend a check on the behaviour of $\bar{h}_{11}$ for different data size T.

In sum, for a parameter, θ_1, and its posterior precision, $\bar{h}_{11}$:

$$\lim_{T\to\infty} T^{-1}\bar{h}_{11} = \begin{cases} 0 & (\bar{h}_{11} \text{ improves at rate slower than } T) \quad \text{if unidentified} \\ \text{a number} & (\bar{h}_{11} \text{ improves at rate } T) \quad \text{if identified} \end{cases}.$$

Putting it in a simpler form, $\bar{H}_T$ can be inverted to obtain following diagnostic value:

$$T^{-1}\bar{h}_{ii} = T^{-1}\tilde{H}_{ii}^{-1} \quad \text{where } \tilde{H}_{ii} \text{ is the } i\text{-th diagonal element of } \bar{H}_T^{-1}. \quad (4)$$

Although the covariance structure provides some important information, our baseline task is to check the reciprocal of the diagonal elements of $\tilde{H}_T$ for different data size T, where $\tilde{H}_T$ is the inverse of the posterior Hessian, $\bar{H}_T$. More specifically, we check if $\tilde{H}_{ii}^{-1}$ increases at rate T. Alternatively, we can use variances computed from the entire posterior density, say, by using the Markov Chain Monte Carlo (MCMC) Method. Since the Hessian shows the asymptotic precision, which is the inverse of the variance, using

[5] The number subscripts indicate submatrices: e.g., $\bar{H}_{22}$ is $\bar{H}_T$ eliminating its first row and first column. To avoid overly messy notation, we omit subscript T to show data size T, when we discuss submatrices.

the Hessian or (exact) posterior variance are almost equivalent for a large T (though not exactly equivalent for finite T). However, in practice we do not need additional computation to obtain $\bar{H}_T$, as almost all gradient-based maximizing algorithms compute it automatically,[6] while the use of the MCMC typically requires additional computation, which is itself often time consuming.

Note that the KPS indicator focuses on identification by the data as it effectively excludes help by any chosen prior, which is dwarfed as $T \to \infty$. This feature is distinct from other existing diagnostics, in which most cases data is either irrelevant or considered jointly with the prior, and forms a strong motivation in our view for the applied researcher to use KPS.

2.2 Design of Experiment

We investigate the extent to which the key parameters of the Smets and Wouters' (2007) US model are identified. The model equations are listed in Table 1 and the priors and definitions of parameters are presented by Table 2. Our baseline exercise is as follows:

1. Given estimated parameters $\bar{\theta}$, we simulate the model to generate artificial data for, say, $10,000$ periods ($T = 10,000$);

2. We re-estimate the model with $T = 10$, 100, $1,000$ and $10,000$. Every larger sample encompasses the previous smaller sample(s);

3. We check the convergence of the posterior variance of each parameter. Specifically, a parameter is said to be identified, if its variance shrinks faster than or at the same rate of the sample size T. In this case $1 \leq \frac{T=n}{T=N} \frac{\sigma^2_{T=n}}{\sigma^2_{T=N}}$, where n is the shorter sample size with, say, $T = 10$, and 100 and N is the largest sample size, in this case $T = 10,000$.

In step 1, the artificial data set is generated by simulating the model to give the time series of output, consumption, investment, hours worked, inflation, the real wage

[6] For low dimension problems, often non-gradient-based algorithms, such as grid search type methods, are much more efficient. However, since the dimension of estimated parameter is typically large (say, more than 3) for typical DSGE estimations, it is rather exceptional to use an algorithm that does not rely on the Hessian.

and the nominal interest rate as in Smets and Wouters (2007). We have used both the MCMC algorithm and the inverse of the Hessian to obtain the posterior variances.[7] We then examine the rate at which the posterior variance falls, normalized by the increase the sample size of the estimates. We use variance, rather than precision, because, given non-normality, it is not an easy task to recover the precision from MCMC exact variance.

Then using the results of this baseline experiment, we impose several restrictions on certain weakly identified parameters. These restrictions allow us to assess how the result is affected, since fixing some weakly (and non-) identified parameters is common econometric practice. For some parameters, we impose *ad hoc* parameter restrictions such as $\iota_p = \iota_w$ and $\xi_p = \xi_w$, which can be regarded as cross parameter restrictions, where we simply assume wage and price share the same degree of indexation and stickiness.

3 Results

Throughout our exercises, following Smets and Wouters (2007), we fix the capital depreciation, δ, the wage markup in steady-state, ϕ_w, the government consumption to output ratio in steady-state, g_y, the Kimball curvature parameter for goods price elasticity, ϵ_p, and the Kimball curvature for wage elasticity, ϵ_w.[8] It is well known that these parameters are not identified; i.e., with these parameters, the maximization algorithms cannot find the posterior mode. In this respect, we can avoid the Catch 22 problem because we know this fact; in general, however, we have to do some trial and error process to find totally unidentified parameters.

[7] There are a couple of further technical notes here. First, in this experiment, we use Dynare; with it, it is easy to compute the KPS indicator. Second, in some preliminary simulations, the maximization algorithms cannot find the maximum posterior points. Often, this cannot be resolved even after trying several different initial values with different maximization algorithms. In this case, we use a different part of the artificially generated data. More practically, in all exercises, we discard the first 10% of the artificial data to eliminate the effects of the initial state. If, however, the Dynare programme cannot find the maximum point of the posterior, we redo all the exercises by discarding the first 10% plus 1 of the artificial data (keeping $T = 10, 100, 1000, 10000$). In our exercise, longer data sets include shorter ones, and we redo all estimations if the algorithm does not converge. One possible concern is that this shows a lack of robustness in our estimations. However, given the nature of the artificial data, the estimation results are almost identical whichever part of the data is used, especially for large T. Although it is not clear why the convergence depends on such a minor difference in the data sets, it seems unlikely that our estimation results are sensitive to this shift in the artificial data.

[8] See Table 2 for the definitions of the symbols, their priors and posterior results.

3.1 Baseline Exercise

We have checked the identification of 36 parameters of this prototypical DSGE model. The first four columns of Table 3 and 4 report the normalized posterior variances of the estimated parameters generated by the MCMC algorithm and the posterior Hessian respectively. First of all, to check whether variance falls more quickly than sample size we compute the ratio of the normalized variances in the right hand columns of each Table, and find that, if we mechanically apply the cut-off point of 1, only 5 parameters are well-identified. These parameters are the trend growth rate, γ, the AR term of government spending shock, ρ_g, the AR term of productivity shock, ρ_a, the AR term of wage mark-up shock, ρ_w and the MA term of wage mark-up shock, ω_w. However, a number of parameters: ρ_π, σ_{qs}, σ_w and σ_g, are close to 1 and could be classified as identified. On the other hand, the worst identified parameters are the inflation coefficient of the monetary policy rule, r_π, the steady state growth rate of inflation, $\bar{\pi}$, and the steady state growth rate of hours worked, $\bar{l}$. At face value, this is a highly problematic result for researchers who wish to estimate DSGE models.

Second, the results from the posterior variance generated by MCMC and from the Hessian are nearly identical. This supports the use of Hessian, because the additional computational burden to obtain the Hessian is effectively zero. Note that we show the results from the inverse of the Hessian for the comparison sake, but we can use the Hessian as a precision matrix in practice (in this case, divide the Hessian by T). Third, the exogenous shock processes tend to be somewhat better identified; this is a rather common finding in most identification literature (see, for example, KPS (2011) and Iskrev and Ratto (2010)). Fourth, our finding about identified or nearly identified parameters are in line with other papers, such as Iskrev (2010a) and Iskrev and Ratto (2010). In our opinion, weakly identified parameters can be classified as follows:

(a) Level parameters that mainly affect the first moments:

The subjective discount factor, β, hours worked in the steady state, $\bar{l}$, inflation rate in the steady state, $\bar{\pi}$, are poorly identified. One exception is γ, the parameter governing output growth trend, which is an outlier in the sense that it is too well identified. Since

the variables in Smets and Wouters (2007) are log-linearised around the steady state level and hence there are no constant terms in their equilibrium equations, the information about the first moment is discarded. As Canova and Sala (2009) pointed out, having constant terms changes the identification in general. Our conjecture is that, if we do not subtract the means from the log-linearized variables and instead add constant terms in the equations, the identification of these parameters could improve significantly.

(b) Monetary policy parameters:

Most coefficients in the Taylor rule are weakly identified: the interest rate weight on inflation, r_π, the interest rate weight on the output gap, r_y, the weight on the change in output gap, $r_{\Delta y}$, and the persistence in interest rates, ρ_r. This is perhaps not surprising because, as discussed extensively in the literature, a simplified Taylor rule that reacts only to inflation often performs as well as the full Taylor rule as in Smets and Wouters (2007). Since inflation and output are highly correlated, it may not matter whether nominal interest rates react to inflation or the output gap.

(c) Price and wage stickiness related parameters:

The probability that price cannot be reset, ξ_p, the degree of price indexation, ι_p, the probability that wage cannot be reset, ξ_w, and the degree of wage indexation, ι_w, are also weakly identified. One possibility is there are strong collinearities between ξ_p and ι_p and between ξ_w and ι_w, as Iskrev (2010b) suggests, but another possibility is those between ξ_p and ξ_w and between ι_p and ι_w, as Canova and Sala (2009) find. In any case, it seems that the nominal rigidity is too densely parameterized or the formulation of nominal rigidity does not capture the data very well.

(d) Other parameters:

The investment adjustment cost, φ, and the elasticity of labour supply, σ_l, are highly colinear pairwise. If we look at the eigenvectors of the Hessian that correspond to the second and third smallest eigenvalues, they are the two dominating members, with the smallest eigenvalue is almost solely related to $\bar{l}$.[9] However, it is not totally clear why these two economically distinct parameters seem to be colinear.

[9]Some preliminary results of the eigenvector analysis are available upon request.

3.2 Applying Restrictions

Using the results of the baseline experiment, we have imposed certain restrictions to the benchmark DSGE model in order to assess whether a significantly greater number of parameters become identified. Specifically, we have considered the following restrictions: we fixed (a) level parameters $\bar{l}$, β, $\bar{\pi}$, (b) monetary policy parameters r_y, $r_{\Delta y}$ and ρ_r, and (d) investment adjustment cost parameter φ at their posterior mean. Furthermore, motivated by Canova and Sala (2009), we set $\iota_p = \iota_w = \iota$ and $\xi_p = \xi_w = \xi$. We estimate r_π, perhaps the most important parameter in the monetary policy rule. These parameter restrictions, of course, reduce the number of free parameters to be estimated. In the similar vein to this, one possible approach to deal with weakly identified parameters is the reduction of parameters by constructing a profile likelihood, in which we represent some parameters as functions of other parameters.[10]

The parameter identification of the restricted model is presented by Table 5. The main findings are as follows. First, as expected, now r_π is fairly well identified, which supports the view that monetary policy parameters are collinear, perhaps because of the high correlation between output gap and inflation in the data. Second, the indicators of ι and ξ do not improve very much; while the speed of precision improvement of ι is slightly higher than ι_p and ι_w, that of ξ is somewhere between ξ_p and ξ_w. Third, there is a slightly positive effect on other parameters; that is, there is some improvement in the rate at which precision improves, though such a effect is rather small. Fourth, we have checked the second moments and IRFs, but fixing weakly identified parameters changes them only negligibly. This is not surprising because we fix them at their posterior mean in the original estimation. All in all, fixing some weakly identified parameters does not change the model behaviour very much and at the same time it does not so helpful to the identification of most of the parameters in this model. Figures 1 and 2 respectively show the impulse responses of the model to a monetary policy shock with and without restrictions in place, showing there is no significant difference in the model properties. That said, it is also clear in both cases that we remain at some distance from full

[10] We thank Hashem Pesaran for this suggestion, which we will pursue in future work.

identification.

4 Discussion

In this section, we briefly discuss some additional issues. First, in terms of the choice between the Hessian and the posterior variance, we find that the use of the Hessian is to be preferred. As we have shown in Tables 3 and 4, the results are almost identical but the additional computational burden to obtain the Hessian is almost zero but to obtain the posterior variance we typically have to employ time consuming MCMC re-sampling, which can take several hours in each case. Also, for the comparison sake, we the Hessian is inverted in Table 4, due to the difficulty in computing the MCMC based precision. However, to avoid unnecessary inversion, it may be better to treat the Hessian as the precision; i.e., treat the Hessian without inverting it. This can be particularly important because weak identification implies that the Hessian is near singular (i.e., ill conditioned). In this case, we can check the normalized precision, which is the diagonal elements of the Hessian divided by the sample size T.

Second, the initial sample size for our analysis of identification can alter the results. Tables 3 to 4, suggest that it would seem preferable to use the results of the increase in precision between $T = 100$ and $T = 10,000$ rather than the comparison between $T = 10$ and $T = 10,000$. This is partly because the variance estimates with a sample size of $T = 10$ seems to be unstable. In fact, if we repeat this experiment several times, we have fairly consistent asymptotic variance estimation for $T = 100$ but it fluctuates to a considerable degree for $T = 10$. This observation is hardly surprising as the estimation of variance is not likely to be well determined with such a short data span. More importantly, it seems that there is a systematic bias for $T = 10$; that is, since, for $T = 10$, the impact of the prior is stronger than when $T = 100$. However, as we discussed, one of the distinct features of the KPS indicator is that it solely focuses on the identification by data (i.e., identification without relying the prior), but the effect of the precision of the prior is not negligible for small T. Hence, the precision for $T = 10$

can be too high because of the prior and as a result the improvement appears to be slow. In our experiment, however, as is clear in Tables 3 and 4, the relative order of identifiability does not change very much when we move from $T = 10$ to $T = 10,000$.[11]

Third, we might be wary in applying a mechanical cut-off rule. We claim that, if a parameter exhibits a precision improvement greater that 1, it is perhaps safe to judge it is well identified. However, even if its speed is slightly lower than 1, it may be still well identified. The KPS method is not a test but an indicator, so we must be cautious in its application, as it is possible that it may sentence too many culprits.

Fourth, these observations lead us to conjecture that it may ultimately be better if we do the same experiment with the data likelihood, rather than with the posterior density. However, separating the data likelihood from the posterior density is quite difficult, and this may not be practical. Accordingly, we leave this exercise to future work.

5 Conclusions

While several identification indicators have been developed for DSGE models, the KPS method is highly attractive in the sense that only it focuses on the data identification, i.e. identification without the help of the prior. There may be some use in combining KPS with other methods, for example, Iskrev's (2010a) J_2, which relies only on the model structure without referring to the data availability. Hence combining these distinct indicators helps us to detect the source of the identification failure. For example, if a parameter of a model passes the J_2 criterion but not the KPS, then we know such an identification problem is because of the lack of sufficient data. In addition, like other Hessian based indicators, the KPS method is also subject to the Catch 22 problem; without *a priori* knowledge about the parameters that are perfectly unidentified, some trial and error may be required to obtain the likelihood (or posterior) mode. In this respect, again, it may be wise to combine it with Jacobian based methods, which often do not rely on the data.

[11] We note a similar improvement in the ratio for $T = 1,000$ vs. $T = 10,000$. However, the relative order does not change much.

In our simple experiments, we find that many parameters in the Smets and Wouters (2007) model, which now works as a benchmark in many DSGE applications, are weakly identified: especially, parameters related to (a) level, (b) monetary policy rule and (c) price and wage stickiness. These findings are rather similar to those in the emerging literature and are also clearly demonstrated by the KPS measure of posterior precision.

In practice, we recommend to using the Hessian (rather than the posterior variance) in KPS method, because of the computational consideration. Also, it may be better to check the change between $T = 100$ and $T = 10,000$, rather than that between $T = 10$ and $T = 10,000$. Finally, given the tendency in KPS, even if a parameter exhibits a precision improvement slower than the order that is theoretically suggested, mechanically judging it as poorly identified may not be the best strategy, as some restrictions may be brought to bear from economic theory to aid identification. To conclude a parameter is poorly identified, its speed of precision improvement must be low and stubbornly so with respect to various model restrictions. That said, the simplicity of the KPS indicator and the extent to which such a widely used workhorse model seems less than fully identified must form a concern for those using Bayesian estimation techniques on DSGE models.

References

An, S. and F. Schorfheide (2005). "Bayesian Analysis of DSGE Models." CEPR Discussion Papers, No. 5207.

Andrle, M. (2010). "A Note on Identification Patterns in DSGE Models." Working Paper Series, No. 1235, European Central Bank.

Canova, F. and L. Sala (2009). "Back to Square One: Identification Issues in DSGE Models." Working Paper Series, No. 583, European Central Bank.

Consolo, A., C. A. Favero, and A. Paccagnini (2007). "On the Statistical Identification of DSGE Models." Working Papers 324, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

Iskrev, N. (2008). "Evaluating the Information Matrix in Linearized DSGE Models." Economics Letters, No. 99, pp607-610.

Iskrev, N. (2010a). "Local Identification in DSGE Models." Journal of Monetary Economics, No. 57, pp189-202.

Iskrev, N. (2010b). "Parameter Identification in Dynamic Economic Models." Economic Bulletin and Financial Stability Report Articles. Banco de Portugal, Economics and Research Department.

Iskrev, N. and M. Ratto (2010). "Analysing Identification Issues in DSGE Models." Mimeo.

Komunjer, I. and S. Ng (2011). "Dynamic Identification of DSGE Models." Mimeo.

Koop, G., M. H. Pesaran, and R. P. Smith (2011). "On Identification of Bayesian DSGE Models." CESifo Working Paper Series, No. 3423, CESifo Group Munich.

Ratto, M., S. Tarantola, A. Saltelli and P. C. Young (2004). "Accelerated Estimation of Sensitivity Indices Using State Dependent Parameter Models.", Mimeo.

Ratto, M. (2008). "Analysing DSGE Models with Global Sensitivity Analysis." Computational Economics, No. 31, pp115-139.

Ratto, M., W. Roeger, and J. Veld (2009). "Quest III: An Estimated Open-Economy DSGE Model of the Euro Area with Fiscal and Monetary Policy." Economic Modelling, No. 26, pp222-233.

Sims, Christopher A, (2002). "Solving Linear Rational Expectations Models," Computational Economics, Springer, vol. 20(1-2), pages 1-20, October.

Smets, F. and R. Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." CEPR Discussion Papers, No. 6112.

Table 1: Log-linearized equations of the DSGE model of Smets and Wouters (2007)

(1) $y_t = c_y c_t + i_y i_t + z_y z_t + \epsilon_t^g$

(2) $c_t = \frac{h/\gamma}{1+h/\gamma} c_{t-1} + (1 - \frac{h/\gamma}{1+h/\gamma}) E_t c_{t+1} - \frac{(\sigma_c - 1)(WL/C)}{\sigma_c (1+h/\gamma)} (l_t - E_t l_{t+1}) + \frac{(1-h/\gamma)}{(1+h/\gamma)\sigma_c} (r_t - E_t \pi_{t+1} + \epsilon_t^b)$

(3) $i_t = \frac{1}{1+\beta\gamma^{1-\sigma_c}} i_{t-1} + (1 - \frac{1}{1+\beta\gamma^{1-\sigma_c}}) E_t i_{t+1} + \frac{1}{(1+\beta\gamma^{1-\sigma_c})\gamma^2\varphi} q_t + \epsilon_t^i$

(4) $q_t = \frac{1-\delta}{Rk+(1-\delta)} E_t q_{t+1} + (1 - \frac{1-\delta}{Rk+(1-\delta)}) r_{t+1}^k - (r_t - \pi_{t+1} + \epsilon_t^b)$

(5) $k_t = \frac{1-\delta}{\gamma} k_{t-1} + (1 - \frac{1-\delta}{\gamma}) i_t + (1 - \frac{1-\delta}{\gamma})(1 + \beta\gamma^{1-\sigma_c})\gamma^2\varphi\epsilon_t^i$

(6) $k_t^s = k_{t-1} + z_t$

(7) $z_t = \frac{1-\Psi}{\Psi} r_t^k$

(8) $r_t^k = -(k_t - l_t) + w_t$

(9) $y_t = \phi(\alpha k_t^s + (1-\alpha) l_t + \epsilon_t^a)$

(10) $\pi_t = \frac{\iota_p}{(1+\beta\gamma^{1-\sigma_c})\iota_p} \pi_{t-1} + \frac{\beta\gamma^{1-\sigma_c}}{1+\beta\gamma^{1-\sigma_c}\iota_p} E_t \pi_{t+1} + \frac{1}{(1+\beta\gamma^{1-\sigma_c})\iota_p} \frac{(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)}{((\phi_p - 1)\epsilon_p + 1)\xi_p} \mu_t^p + \epsilon_t^p$

(11) $\mu_t^p = \alpha(k_t^s - l_t) - w_t + \epsilon_t^a$

(12) $w_t = \frac{1}{1+\beta\gamma^{1-\sigma_c}} w_{t-1} + (1 - \frac{1}{1+\beta\gamma^{1-\sigma_c}})(E_t w_{t+1} + E_t \pi_{t+1}) - \frac{1+\beta\gamma^{1-\sigma_c}\iota_w}{1+\beta\gamma^{1-\sigma_c}} \pi_t + \frac{\iota_w}{1+\beta\gamma^{1-\sigma_c}} \pi_{t-1} - \frac{1}{(1+\beta\gamma^{1-\sigma_c})\iota_w} \frac{(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)}{((\phi_w - 1)\epsilon_w + 1)\xi_w} \mu_t^w + \epsilon_t^w$

(13) $\mu_t^w = w_t - (\sigma_l l_t + \frac{1}{1-\lambda}(c_t - \lambda c_{t-1}))$

(14) $r_t = \rho_r r_{t-1} + (1-\rho)(r_\pi \pi_t + r_Y (y_t - y_t^p)) + r_{\Delta y}[(y_t - y_t^p) + (y_{t-1} - y_{t-1}^p)] + \epsilon_t^R$

(15) $\epsilon_t^a = \rho_a \epsilon_{t-1}^a + \eta_t^a$

(16) $\epsilon_t^g = \rho_g \epsilon_{t-1}^g + \eta_t^g + \rho\eta_t^a$

(17) $\epsilon_t^i = \rho_i \epsilon_{t-1}^i + \eta_t^i$

(18) $\epsilon_t^b = \rho_b \epsilon_{t-1}^b + \eta_t^b$

(19) $\epsilon_t^w = \rho_w \epsilon_{t-1}^w + \eta_t^w + \mu_w \eta_t^w$

(20) $\epsilon_t^p = \rho_p \epsilon_{t-1}^p + \eta_t^p + \mu_p \eta_t^p$

(21) $\epsilon_t^r = \rho_r \epsilon_{t-1}^r + \eta_t^r$

Note:The model has fourteen endogenous variables: y, output, c, consumption, i, investment, q, price of installed capital, k, total capital stock, k^s, the amount of capital used in production, z, capital utilisation rate, r^k, rental rate of capital, π, inflation, w, wages, r, nominal interest rate, μ^w, wage mark up and μ^w, price mark up. And the responses of fourteen endogenous variables are driven by seven shocks: ϵ^a, total factor productivity, ϵ^i, aggregate investment, ϵ^b, consumer spending, ϵ^p, price mark-up, ϵ^w, wage mark-up, and ϵ^r, monetary policy shock. As standard, the key behavioural equations are obtained by deriving optimality conditions for household and firm behaviour. These decision rules are then linearised around their steady-state in standard fashion.This model and the set of exogenous shock processes are estimated on time series data using Dynare.

Table 2: Prior and posterior distributions

		Prior			Posterior		
Par.	Definition	Density	Mean	Std.	Mode	Mean	Std.
φ	Investment adj. cost	N	4.00	1.50	5.47	5.75	1.03
σ_c	Inv. elats. intert. subst.	N	1.50	0.37	1.42	1.38	0.14
h	Consump. habit	B	0.70	0.10	0.73	0.71	0.04
ξ_w	Calvo wage	B	0.50	0.10	0.73	0.70	0.07
σ_l	Elast. labour supply	N	2.00	0.75	1.87	1.77	0.61
ξ_p	Calvo price	B	0.50	0.10	0.65	0.65	0.06
ι_w	Index. of wages	B	0.50	0.15	0.60	0.57	0.13
ι_p	Index. of prices	B	0.50	0.15	0.22	0.25	0.09
Ψ	Capital utilization	B	0.50	0.15	0.54	0.55	0.12
ϕ	Fixed cost	N	1.25	0.12	1.60	1.61	0.08
r_π	Response to inflation	N	1.50	0.25	2.02	2.04	0.18
ρ_r	Interest rate smooth.	N	0.75	0.10	0.81	0.81	0.02
r_y	Response to output	N	0.12	0.05	0.09	0.09	0.02
$r_{\Delta y}$	Response to outp. gap	N	0.12	0.05	0.22	0.23	0.03
$\overline{\pi}$	SS inflation	G	0.62	0.10	0.76	0.78	0.11
$100(\beta^{-1}-1)$	Discount factor	G	0.25	0.10	0.14	0.17	0.06
$\bar{l}$	SS hours worked	N	0	2.00	0.72	0.63	1.07
$100(\gamma-1)$	Trend growth	N	0.40	0.10	0.43	0.43	0.01
α	Share of capital	N	0.30	0.05	0.19	0.19	0.02
δ	Depreciation rate	n.a.	0.025	n.a.	n.a.	n.a.	n.a.
g_y	Government/Output	n.a.	0.18	n.a.	n.a.	n.a.	n.a.
ϕ_w	Wage mark-up	n.a.	1.5	n.a.	n.a.	n.a.	n.a.
ϵ_w	Kimball (wage)	n.a.	10	n.a.	n.a.	n.a.	n.a.
ϵ_p	Kimball (price)	n.a.	10	n.a.	n.a.	n.a.	n.a.
ρ_a	AR prod. shock	Beta	0.50	0.20	0.96	0.96	0.01
ρ_b	AR risk premium	Beta	0.50	0.20	0.18	0.23	0.08
ρ_g	AR government spend.	Beta	0.50	0.20	0.98	0.98	0.01
ρ_{qs}	AR invest. demand	Beta	0.50	0.20	0.70	0.71	0.06
ρ_{ms}	AR monetary policy	Beta	0.50	0.20	0.12	0.14	0.06
ρ_π	AR price mark-up	Beta	0.50	0.20	0.91	0.14	0.05
ρ_w	AR wage mark-up	Beta	0.50	0.20	0.97	0.97	0.01
ω_p	MA price mark-up	Normal	0.50	0.20	0.74	0.72	0.09
ω_g	Prod. shock in G	Normal	0.50	0.25	0.52	0.52	0.09
ω_w	MA wage mark-up	Normal	0.50	0.20	0.89	0.85	0.05
σ_a	Std. prod. shock	IG	0.10	2.00	0.45	0.46	0.03
σ_b	Std. risk premium	IG	0.10	2.00	0.24	0.24	0.02
σ_g	Std. government	IG	0.10	2.00	0.52	0.53	0.03
σ_{qs}	Std. investment	IG	0.10	2.00	0.45	0.45	0.05
σ_{ms}	Std. mon. pol.	IG	0.10	2.00	0.24	0.24	0.01
σ_π	Std. price mark-up	IG	0.10	2.00	0.14	0.14	0.02
σ_w	Std. wage mark-up	IG	0.10	2.00	0.25	0.24	0.02

Table 3: Normalized posterior variances of structural parameters (generated by MCMC methods)

	Normalized Variance				Identification Ratio			
Parameter	T=10	T=100	T=1000	T=10,000	T=10/T=10,000		T=100/T=10,000	
φ	15.53	78.46	210.3	415.4	γ	19507.20	γ	8694.97
σ_c	0.165	0.168	0.398	0.782	ρ_g	41.72	ρ_g	24.36
h	0.922	1.828	3.114	4.900	ρ_a	26.27	ρ_a	9.44
ξ_w	0.159	0.871	1.594	1.884	ρ_w	10.22	ρ_w	5.19
σ_l	4.20	40.74	137.5	221.4	ω_w	1.31	ω_w	4.15
ξ_p	0.029	0.196	0.349	0.320	ρ_π	0.82	α	1.40
ι_w	0.009	0.047	0.064	0.072	σ_{qs}	0.82	σ_{qs}	1.22
ι_p	0.131	0.562	1.194	1.587	σ_w	0.78	ρ_π	1.08
Ψ	0.056	0.208	0.379	0.712	σ_g	0.74	σ_w	1.01
ϕ	0.209	1.658	3.932	4.438	ρ_{qs}	0.39	ρ_{qs}	0.80
r_π	0.536	3.055	18.35	53.79	ρ_b	0.32	σ_g	0.79
ρ_r	0.183	1.272	2.035	2.416	ω_g	0.30	ρ_b	0.77
r_y	0.022	0.089	0.250	0.552	σ_{ms}	0.29	σ_b	0.71
$r_{\Delta y}$	0.096	0.918	2.556	5.980	σ_a	0.27	ι_w	0.65
$\overline{\pi}$	4.63	58.40	309.5	557.8	σ_c	0.21	σ_{ms}	0.65
β	0.024	0.142	0.696	1.471	α	0.20	σ_a	0.62
$\bar{l}$	0.101	0.755	7.437	14.16	h	0.19	ξ_p	0.61
γ	0.072	0.032	0.000	0.000	σ_b	0.18	ω_g	0.59
α	0.208	1.430	0.965	1.023	ω_p	0.16	ρ_r	0.53
ρ_a	0.553	0.199	0.022	0.021	ι_w	0.12	ξ_w	0.46
ρ_b	0.055	0.132	0.128	0.172	σ_π	0.12	ρ_{ms}	0.43
ρ_g	0.493	0.288	0.020	0.012	ρ_{ms}	0.10	ϕ	0.37
ρ_{qs}	0.043	0.088	0.094	0.110	ξ_p	0.09	h	0.37
ρ_{ms}	0.119	0.534	0.984	1.251	ξ_w	0.08	ι_p	0.35
ρ_π	0.310	0.407	0.319	0.378	ι_p	0.08	ω_p	0.31
ρ_w	0.292	0.148	0.018	0.029	Ψ	0.08	Ψ	0.29
ω_p	0.060	0.118	0.216	0.386	ρ_r	0.08	σ_π	0.25
ω_g	0.017	0.033	0.055	0.056	ϕ	0.05	σ_c	0.22
ω_w	0.433	1.365	0.273	0.329	r_y	0.04	φ	0.19
σ_a	0.406	0.927	1.346	1.493	φ	0.04	σ_l	0.18
σ_b	0.431	1.699	2.038	2.386	σ_l	0.02	r_y	0.16
σ_g	0.115	0.122	0.149	0.154	β	0.02	$r_{\Delta y}$	0.15
σ_{qs}	0.305	0.452	0.401	0.372	$r_{\Delta y}$	0.02	$\overline{\pi}$	0.10
σ_{ms}	0.172	0.384	0.511	0.594	r_π	0.01	β	0.10
σ_π	0.052	0.109	0.237	0.429	$\overline{\pi}$	0.01	r_π	0.06
σ_w	0.078	0.101	0.089	0.100	$\bar{l}$	0.01	$\bar{l}$	0.05

Notes for Table 3-5: Normalized variance is the estimated variance times T. And the Identification Ratio shows the ratio of the Normalized variance for different T. A ratio greater than 1 shows that convergence is faster than T.

Table 4: Normalized posterior variances of structural parameters (generated by H^{-1} method)

	Normalized Variance				Identification Ratio			
Parameter	T=10	T=100	T=1000	T=10,000	T=10/T=10,000		T=100/T=10,000	
φ	18.12	85.26	232.9	428.8	γ	18360.76	γ	7052.12
σ_c	0.813	1.911	3.342	5.256	ρ_g	149.03	ρ_g	30.30
h	0.072	0.116	0.220	0.432	ρ_a	6.49	ρ_a	9.50
ξ_w	0.042	0.202	0.435	0.727	ρ_w	6.43	ρ_w	3.74
σ_l	4.147	41.98	154.61	220.0	ω_w	2.07	ω_w	3.48
ξ_p	0.026	0.260	0.329	0.324	ρ_π	0.68	ρ_b	1.51
ι_w	0.289	1.744	3.793	4.686	σ_{qs}	0.61	σ_w	0.95
ι_p	0.109	0.866	1.637	1.826	σ_w	0.39	σ_{qs}	0.88
Ψ	0.263	1.533	1.909	2.248	σ_g	0.35	σ_g	0.87
ϕ	0.107	0.564	1.454	1.688	ω_p	0.34	σ_a	0.84
r_π	0.003	0.043	0.062	0.067	ω_g	0.34	σ_b	0.84
ρ_r	0.060	0.152	0.246	0.423	ρ_b	0.30	ξ_p	0.80
r_y	0.024	0.080	0.254	0.554	σ_b	0.28	ω_g	0.73
$r_{\Delta y}$	0.024	0.168	0.631	1.630	α	0.27	ρ_π	0.69
$\overline{\pi}$	0.074	0.856	7.898	14.52	ρ_{qs}	0.25	Ψ	0.68
β	0.100	0.783	2.876	5.848	σ_a	0.17	σ_π	0.64
$\bar{l}$	1.474	53.04	327.2	507.6	σ_{ms}	0.17	ρ_{qs}	0.59
γ	0.070	0.027	0.000	0.000	h	0.17	α	0.57
α	0.016	0.034	0.054	0.060	σ_c	0.15	ω_p	0.49
ρ_a	0.150	0.220	0.024	0.023	ρ_r	0.14	ι_p	0.47
ρ_b	0.296	1.516	0.983	1.001	Ψ	0.12	ρ_{ms}	0.47
ρ_g	1.839	0.374	0.018	0.012	ρ_{ms}	0.11	ι_w	0.37
ρ_{qs}	0.154	0.370	0.563	0.628	ξ_p	0.08	σ_c	0.36
ρ_{ms}	0.146	0.641	0.927	1.374	ϕ	0.06	ρ_r	0.36
ρ_π	0.267	0.273	0.372	0.395	ι_w	0.06	ϕ	0.33
ρ_w	0.168	0.098	0.019	0.026	ι_p	0.06	ξ_w	0.28
ω_p	0.830	1.193	1.985	2.429	ξ_w	0.06	h	0.27
ω_g	0.508	1.084	1.368	1.491	σ_π	0.04	σ_{ms}	0.21
ω_w	0.641	1.078	0.269	0.310	r_y	0.04	φ	0.20
σ_a	0.026	0.125	0.147	0.148	φ	0.04	σ_l	0.19
σ_b	0.031	0.091	0.100	0.109	σ_l	0.02	r_y	0.15
σ_g	0.056	0.140	0.155	0.161	β	0.02	β	0.13
σ_{qs}	0.244	0.352	0.342	0.402	$r_{\Delta y}$	0.01	$\bar{l}$	0.10
σ_{ms}	0.144	0.180	0.422	0.853	r_π	0.01	$r_{\Delta y}$	0.10
σ_π	0.003	0.043	0.062	0.067	$\overline{\pi}$	0.01	r_π	0.06
σ_w	0.039	0.096	0.090	0.100	$\bar{l}$	0.00	$\overline{\pi}$	0.06

Table 5: Normalized posterior variances of the restricted model

	Normalized Variance				Identification Ratio			
Parameter	T=10	T=100	T=1000	T=10,000	T=10/T=10,000		T=100/T=10,000	
σ_c	0.941	1.264	2.423	2.634	γ	14773.96	γ	7939.46
h	0.059	0.084	0.137	0.145	ρ_g	130.16	ρ_g	28.64
σ_l	4.865	35.227	51.889	47.222	ρ_w	8.28	σ_l	4.60
ξ_p	0.017	0.132	0.206	0.217	ρ_a	2.86	h	2.74
ι_p	0.134	0.795	1.189	1.253	ω_w	0.95	ϕ	2.61
Ψ	0.274	1.391	1.834	1.953	σ_{ms}	0.69	ρ_a	1.63
ϕ	0.097	0.622	1.316	1.415	σ_{qs}	0.64	ω_g	1.34
r_π	0.224	0.698	0.471	0.428	α	0.61	σ_{qs}	1.07
γ	0.054	0.029	0.000	0.000	ρ_π	0.52	σ_b	1.07
α	0.028	0.031	0.046	0.045	ω_p	0.49	α	1.01
ρ_a	0.059	0.096	0.022	0.021	h	0.41	σ_g	0.95
ρ_b	0.390	1.047	0.963	0.977	ρ_b	0.40	ρ_b	0.92
ρ_g	1.418	0.312	0.017	0.011	σ_g	0.39	σ_w	0.92
ρ_{qs}	0.151	0.336	0.485	0.499	σ_c	0.36	σ_c	0.75
ρ_{ms}	0.083	0.357	0.471	0.500	ω_g	0.35	ω_p	0.71
ρ_π	0.224	0.698	0.471	0.428	σ_b	0.33	σ_a	0.71
ρ_w	0.195	0.062	0.015	0.024	ρ_{qs}	0.30	ρ_{ms}	0.71
ω_p	0.687	1.860	1.422	1.391	σ_w	0.30	ρ_{qs}	0.70
ω_g	0.502	1.029	1.333	1.447	ρ_{ms}	0.17	σ_{ms}	0.69
ω_w	0.333	0.964	0.290	0.351	σ_a	0.16	σ_π	0.67
σ_a	0.022	0.144	0.146	0.143	Ψ	0.14	Ψ	0.65
σ_b	0.032	0.065	0.094	0.099	r_π	0.11	ξ_p	0.63
σ_g	0.061	0.144	0.153	0.156	ι_p	0.11	ω_w	0.61
σ_{qs}	0.229	0.339	0.332	0.356	σ_l	0.10	ρ_π	0.58
σ_{ms}	0.161	0.162	0.224	0.233	ξ_p	0.08	ρ_w	0.48
σ_π	0.003	0.038	0.042	0.042	ϕ	0.07	ι_p	0.48
σ_w	0.022	0.078	0.068	0.073	σ_π	0.06	r_π	0.44

Note: Variances are generated by H^{-1} method. Using the results shown by Table 3 and 4, following parameters are fixed at their posterior means and not estimated: unidentified two steady state growth parameters, $\overline{\pi}$ and $\bar{l}$, three parameters of monetary policy reaction function, ρ_r , r_y, $r_{\Delta y}$, Also the two wage parameters ξ_w and ι_w are set so that ξ_w=ξ_p and ι_w= ι_p.

Table 6: Analysis of second moments and correlations with output

	Data			Non-estim. model			Unrest. model			Rest. model		
Variable	Std.	R. std.	Corr.	Std.	R. std.	Corr.	Std.	R. std.	Corr.	Std	R. std.	Corr.
Output	0.87	1.00	1.00	0.45	1.00	1.00	0.87	1.00	1.00	0.87	1.00	1.00
Consumption	0.70	0.80	0.67	0.41	0.91	0.93	0.63	0.72	0.60	0.61	0.71	0.59
Investment	2.25	2.59	0.67	0.49	1.09	0.85	2.14	2.46	0.59	2.15	2.49	0.59
Real wage	0.56	0.64	0.09	0.26	0.58	0.75	0.51	0.59	0.22	0.53	0.62	0.23
Hours Worked	2.91	3.34	0.13	0.55	1.25	0.31	1.03	1.18	0.17	0.99	1.15	0.17
Inflation	0.61	0.70	-0.32	0.27	0.61	0.14	0.36	0.41	-0.22	0.33	0.38	-0.19
Nominal interest	0.83	0.95	-0.25	0.23	0.52	-0.16	0.37	0.43	-0.23	0.36	0.42	-0.21

Note: Above moments are computed using artificial data generated by estimated DSGE models apart from non-estimated model. Model economy simulated for 12000 periods for all shocks, first 1000 omitted and the remaining is used for computing the moments and correlations.

Figure 1: The responses of key model variables to orthogonalized nominal interest rate shock: Unrestricted model

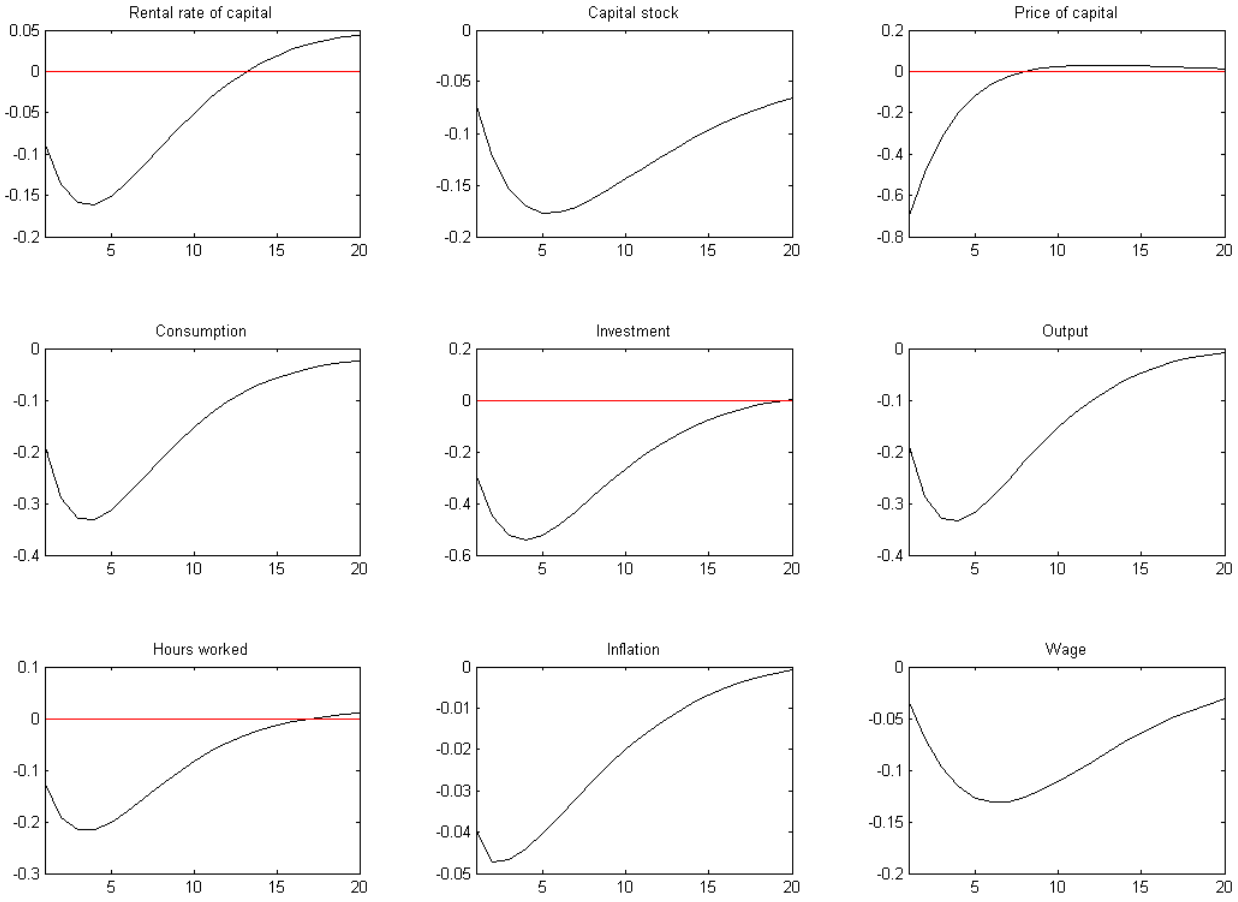

Figure 2: The responses of key model variables to orthogonalized nominal interest rate shock: Restricted model