Bayesian Estimation of DSGE models: Identification using a diagnostic indicator*

Jagjit S. Chadha† and Katsuyuki Shibayama‡

NIESR and University of Kent

September 2018

Abstract

Koop, Pesaran and Smith (2013) 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. As it does not require any additional programming, a researcher just needs to generate artificial data and estimate the model with increasing sample size, T. We apply this indicator to the benchmark Smets and Wouters' (2007) DSGE model of the US economy, and suggest how to implement this indicator on DSGE models.

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

JEL CLASSIFICATION: C51, C52, E32.

*__Acknowledgements__: We thank Herman Dijk, Gary Koop, Chris Otrok, Keisuke Otsu, Hashem Pesaran and Ron Smith for advice and support, as well as an anonymous referee. This paper is forthcoming in the Journal of Economics Dynamics and Control.

†National Institute of Economic and Social Research and Centre for Macroeconomics. E-mail: j.chadha@niesr.ac.uk.

‡Department of Economics, University of Kent, Canterbury. E-mail: k.shibayama@kent.ac.uk

1 Introduction

Many macroeconomists have expressed concern about the extent to which identification of structural, or DSGE, models may or may not have been achieved during estimation.[1] Reflecting the rapid progress of Bayesian estimation techniques, it is now common practice to estimate DSGE models rather 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, DSGE models involve nonlinear estimation with many theoretical parameter restrictions and accordingly identification may be considerably more problematic.

And yet it gets worse: in the Bayesian framework the prior often masks the problem of non- or weak identification by the data.[2] 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.* (2013) discuss, from a pure Bayesian perspective, that 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 position 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

[1] DSGE stands for dynamic stochastic general equilibrium and refers to structural models derived from microfoundations, which are perturbed by a rich structure of shocks.

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

indicators has been pioneered by Iskrev (2010a), Iskrev and Ratto (2010) and Komunjer and Ng (2011). Typically, this intermediate target is a set of data moments. If the 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.[3] Note that this type of diagnostic is only a necessary condition for identification; in the sense that even if a proper Jacobian is column full rank, quite often the limitation of data availability hampers parameter estimation. This circumstance however implies that by combining a Jacobian-based and a Hessian-based approach, which we discuss in the next paragraph, we can detect the source of non- or weak identification. For example, if a Jacobian is column full rank but a Hessian is not full rank, then we could conclude the failure of identification is not because of the model structure but because of data limitations.

The second line of indicators, such as Koop *et al.* (2013, 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 nearly all maximizing

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

algorithms 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.[4] Importantly this means that this class of indicators work only for weakly identified parameters; a researcher has to obtain *a priori* information about the parameters that are 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.[5]

The purpose of our paper 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 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. As we shall explain, all a researcher then has to do, when estimating any model, is simply to check the speed at which the parameter precision increases. On the basis of our results, we will recommend using an Identification Ratio that compares the estimates with a sample of either 1,000 or 5,000 observations with those of 10,000.

As the second method is more widely applicable, it is the one that we apply in the analysis of the identification of the influential DSGE model of Smets and Wouters (2007) and discuss several practical issues in computing and interpreting the simple KPS

[4]Even for Jacobian based-methods, however, we often need data to pin down the point in the parameter space, at which we calculate a proper Jacobian matrix. If so, there is a similar sort of Catch-22 problem even for Jacobian based-methods.

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

indicator.[6] The results are clear enough to allow us to make a number of observations. As many researchers use Smets and Wouters (SW), or its variants, as a testing ground for their identification methods, we are thus able to match our results in using a very simple indicator with theirs, see, for example, Iskrev (2010a) and Iskrev and Ratto (2010). Although we will need to investigate other key models, as well, to be conclusive, broadly speaking, because our findings on the SW model are consistent with other results, we should continue to be cautious about whether estimated parameters are indeed identified. The issues on identification in such a widely cited model, suggest that there continues to be a question mark about whether Bayesian estimation of DSGE models generates more heat than light. And so what we can suggest is that we should accordingly use a simple indicator to examine identification. In our view a little more clarity about identification when using DSGE procedures would aid and focus the debate on the development of models with more realistic economic structures. The regular use of KPS when estimating any DSGE model would allow the reader to make up their own mind on the question of whether model identification has been achieved of the estimates presented.

The rest of the paper is organized as follows: Section 2 briefly introduces the idea of 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 outlining the intuition of the KPS indicator.[7] Consider the Bayesian estimation of a DSGE model. Let $\theta = (\theta_1, \theta_2 ... \theta_n)$ be a parameter vector, T be the size of the data and note that an underscore refers to the prior, an overscore as the posterior and a circumflex refers to an estimated parameter. Suppose that the posterior

[6] According to Google Scholar (accessed in June 2012) the SW paper and its variants has 1,269 citations.

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

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, $\hat{S}_T$ (which measures data precision, $T\hat{S}_T$, divided by T) converges to a certain point, as 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, the parameter mean is $\bar{\theta}_{1T}$ and the precision of the estimate, $\bar{h}_{11}$, is given as $\bar{h}_{11} = \bar{H}_{11} - \bar{H}_{12}\bar{H}_{22}^{-1}\bar{H}_{21}$.[8] Hence, we obtain

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

Following the same analysis as above, at the limit this expression (3) is given as,

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

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 standard, or frequentist, econometric theory, it is easy

[8] The numerical 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.

to see, if $\bar{\theta}_1$ is well identified, $\hat{S}_{11}^{-1}$ approaches a particular 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 observation means that, as the data sample become larger, the estimation will become more precise. These observations lead KPS to recommend checking the behavior of $\bar{h}_{11}$ for increasingly large data sizes, T.

In sum, for a given 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 the following diagnostic value:

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

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 increasingly large data sizes 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 that the asymptotic precision, which is the inverse of the variance, using the Hessian or (exact) posterior variances are almost equivalent for a large T (though not exactly equivalent for smaller T). However, in practice we do not need additional computation to obtain $\bar{H}_T$, as almost all gradient-based maximizing algorithms compute the posterior Hessian automatically,[9] 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; where in most cases data is either irrelevant or considered

[9]For problems of low dimensionality non-gradient based-algorithms, such as grid search type methods, are often much more efficient. However, since the dimension of the estimated parameters 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.

jointly with the prior. This feature creates a strong motivation, in our view, for the applied researcher to use KPS. Researchers will though need guidance on three key issues: (i) what values of the sample size, T, ought to be used as the baseline and then the comparator as the sample size increases; (ii) whether we ought to concentrate on the MCMC chain or the Hessian to calculate variances; and (iii) given that precision increases with sample size, at what rate of improvement should we consider the model identified? As a result of our work, we are in a position to suggest how to use the KPS indicator for applied analysis.

2.2 Design of Experiments

We investigate the extent to which the key parameters of the benchmark Smets and Wouters' (2007) model of the US macroeconomy are identified. This widely-cited model can be thought of as a standard depiction of a key aggregate relationships and has been widely used in the macroeconometric literature. For clarity, the model equations are listed in Table 1 and the priors, posterior modes and definitions of parameters are presented in 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, $1,000$, $5,000$ and $10,000$. Note that every larger sample encompasses the previous smaller sample(s);

3. We check the convergence of the posterior variance of each parameter. We define identification ratio as $IR_n = n\sigma^2_{T=n} / \left(N\sigma^2_{T=N}\right)$, where n is the shorter sample size with, say, $T = 1,000$, and $5,000$, and $N = 10,000$ is the largest sample size.

4. A parameter is said to be identified, if its posterior variance shrinks faster than or at the same rate as the sample size T used in the estimation.[10] That is, if identified, theoretically, $IR_n \geq 1$.

[10] Note that $T^{-1}\bar{h}_{ii}$ converges to a number if a parameter i is identified, the ratio of it for two different sample sizes should go to 1.

The standard estimation of the SW model requires seven macroeconomic time series: output, consumption, investment, hours worked, inflation, the real wage and the nominal interest rate. For Step 1, the artificial data set is generated by simulating the model to give the seven macroeconomic time series. We use both the inverse of the Hessian and the MCMC algorithm to obtain the posterior variances.[11] 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 trivial task to recover the precision from the MCMC measure of exact variance. We primarily examine the magnitude of the convergence statistic for both run these measures of inverse of the Hessian and the MCMC algorithm and consider results for the numerator in intervals from $T = 10$ to $5,000$.

5. We employ some simple restrictions implied by theory to examine the implications for identification of this model.

Once we have established 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.

[11] 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 problem 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.

3 Results

Throughout our results and in order to simulate the artificial data, following Smets and Wouters (2007) and the extant literature, 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.[12] It is well-known that these deep structural parameters are not identified or, at least, very difficult to identify: i.e. for these deep or steady-state parameters, the maximization algorithms cannot find the posterior mode. In this respect, we can avoid a form of Catch 22-type problem because we know this fact from earlier researchers.

3.1 Baseline Exercise

We have checked the identification of 41 parameters of this prototypical structural, or DSGE, model, including the above five parameters eliminated the estimation of Smets and Wouters (2007). The main results are shown in Tables 3 to 6. Tables 3 and 4 correspond to first to the estimated posteriors and then the identification ratio generated by the MCMC and then for Tables 5 and 6 by the Hessian, respectively. The first five columns of Table 3 and 5 report the normalized posterior variances of the estimated parameters generated by the MCMC algorithm and the posterior Hessian respectively.[13] The parameters above the Table-wide horizontal line are structural parameters and the ones below refer to parameters governing the models forcing processes, or shocks.

Our first result is that the results from Table 3 and 5 suggest that the results from the posterior variance generated by MCMC and from the Hessian are nearly identical in parameter and response to increasing the sample size, T. This result strongly supports the use of the Hessian because the additional computational burden to obtain the Hessian is effectively zero while the computational times for the MCMC method for large T takes a considerable computation time. Note that we show the results from the inverse of the

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

[13] The estimation results and code are all available from the authors.

Hessian for Table 5 to facilitate the comparison with Table 3, but we can also simply use the Hessian as a precision matrix in practice, in which case, divide the Hessian by T. We would therefore recommend that the applied researcher simply use the Hessian.

Now we are in a position to check whether the posterior variance falls more quickly than sample size and we simply compute, what we call the Identification Ratio (IR) of the normalized variances in Tables 4 and 6 and find that, if we mechanically apply the cut-off point of 1, (which is indicated by the short horizontal lines in each Table) three issues emerge. First, that a significant number of parameters do not seem to be identified as the IR is under 1. But secondly that there is a tendency for the number of parameters that are identified to increase when the sample size of the numerator is increased. Finally and relatedly, the whole set of ratios tend to move up towards one when we increase the sample size of the numerator. In Table 4 the least well identified parameter and median go from 0.124 to 0.776 and from 0.31 to 0.97 as we move from an IR involving $\frac{T=10}{T=100}$ to one employs $\frac{T=5{,}000}{T=10{,}000}$ and similarly for Table 6 we go from 0.005 to 0.805 and from 0.24 to 1.02, respectively.

So although theoretically the threshold should be 1, many parameters concentrate around 1. This leads us not to suggest the use the theoretical threshold naively. There are mainly two reasons not to use 1 as a threshold: (i) as priors help identification more strongly for small sample than large sample, we may want to judge any tendency to 1 as evidence in support of identification; and (ii) and T even when large but still finite. To illustrate (ii), note that the $IR_{1,000}$ is around 0.1 for some parameters, which is very close to $n/T = 1,000/10,000$, but also far above it for many others. This is not by chance, rather this implies that the precision of these former parameters does not improve at all even if the sample size increases. If we could implement this exercise for $T = \infty$, their RI_∞ should go to 0 as the theory suggests, but not with finite sample size. But for some of the other parameters which show some improvement these observations suggest that we should use the threshold value that should be slightly smaller than 1. Unfortunately, however, it is hard to give a proper threshold *a priori*.

To better illustrate these results, Figures 1 and 2 show the scatter of the IR derived

from the MCMC and the inverse of the Hessian for the two extreme cases of $\frac{T=10}{T=100}$ and $\frac{T=5,000}{T=10,000}$ and whilst we can see that all the points lie close to the implied 45^0 implying little or no difference in the IR from using MCMC or Hessian estimates, the extent of identification is radically different. In Figure 1 nearly all the parameters lie inside the unit box - and thus are not identified - but in Figure 2 nearly all parameters move to cluster around the point of identification.[14]

If we look at the first three columns of Tables 4 and 6, the parameters that are clearly identified 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 parameters that do not seem very well identified 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 but if we move to the IR which employs $\frac{T=1,000}{T=10,000}$ or $\frac{T=5,000}{T=10,000}$ then most of the parameters seem identified or nearly identified. We shall return to this final point.

Overall we also note that the exogenous shock processes tend to be somewhat better identified; this is a rather common finding in most identification literature (see, for example, KPS (2013) and Iskrev and Ratto (2010)). Our findings about identified or nearly identified parameters are in line with other papers, such as Iskrev (2010a) and Iskrev and Ratto (2010). Third, setting aside γ, we can see three groups. The gaps between them appear around 0.3, 0.8 and 1.5. We label these groups as not-identified, weakly identified and strongly identified, and discuss them shortly. For expositional purpose, we look at them on group-by-group basis. Overall, we find the following tendencies: (a) level parameters, (b) monetary policy parameters and (c) nominal stickiness parameters are weakly identified, while (d) parameters for exogenous processes are more or less well identified. Broadly speaking, these findings are not very different

[14] For scaling reasons we do not show any IR estimates larger than 5 but by definition they are clearly identified.

from other papers, such as Iskrev (2010a) and Iskrev and Ratto (2010), although, given difference in the identification designed, our results are not necessarily identical to theirs.

(a) Level Parameters: First, steady state inflation $\bar{\pi}$ and steady state labour hours $\bar{l}$ are only weakly identified. Also, subjective discount factor β is poorly identified, which determines the steady state interest rate. One exception is, as mentioned above, trend growth rate γ, which is identified exceptionally well. They capture the level of the data, but such information is lost in the standard DSGE models, because the endogenous variables are presented as the deviations from their steady state values. 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: Second, monetary policy parameters r_y and r_π are only weakly identified. This could be the problem of data variation, because inflation and output are highly correlated in a Phillips curve type model, the role of r_y and r_π could be similar. This is quite problematic to evaluate the monetary policy performance based on DSGE estimations.

(c) Parameters Related to Nominal Stickiness: Third, the two Kimball parameters for goods and labour aggregators are not identified. This finding is consistent with Iskrev (2010b). Kimball aggregator is an extension of Dixit-Stiglitz aggregator, in which the elasticity of substitution among types of (intermediate) goods is constant. In Kimball aggregator, it is not a constant any more, and the Kimball parameter governs the rate of change in the elasticity of substitution to the change in demand. We can find an intuitive exposition in Iskrev (2010b) but we also would like to point out algebraically that each Kimball parameter appears only together with corresponding Calvo parameter, which shows one minus the price (or wage) reset probability. For example, Kimball curvature parameter ϵ_p for goods price elasticity is coupled with Calvo parameter ξ_p

both appear only in the last term of equation (10) in Table 1. Following Smets and Wouters (2007), we labels the coefficient on that term π_3. If we ignore prior, even if data successfully the value of π_3, there are infinitely many combinations of ϵ_p and ξ_p that attain such a value for π_3. In this sense, non-identification of Kimball parameters is not surprising but it is still a puzzle why two Calvo parameters are identified, though only weakly.

Also, from our previous exercise (results not shown here), there is some evidence to consider that Calvo parameters and indexation parameters ι_p and ι_w are collinear, which also could reduce their identifiability. Indeed, Iskrev (2010b) suggests that there are strong collinearities between ξ_p and ι_p and between ξ_w and ι_w, but another possibility is those between ξ_p and ξ_w and between ι_p and ι_w, as Canova and Sala (2009) find.

(d) Parameters of Exogenous Processes As a general tendency, we also note that the parameters of the exogenous shock processes are somewhat better identified; this is a rather common finding in most identification literature. See, for example, KPS (2013) and Iskrev and Ratto (2010).

(e) Other Parameters Finally, in terms of the parameters excluded in Smets and Wouters (2007), the wage markup in steady-state ϕ_w is actually only weakly identified, but capital depreciation rate δ and the share of the government expenditure g_y in the goods market clearing condition are well-identified. Also, as mentioned above, two Kimball parameters are totally unidentified. Presumably g_y plays a similar role to the standard deviation of the government shock σ_g, through steady state consumption and investment levels, g_y is identifiable. Our result suggests that, for example, if we have more observations, at least potentially g_y and ϕ_w can be estimable in the standard DSGE models.

3.2 Applying Simple Restrictions

Using the results of the baseline experiment, we have imposed a number of simple restrictions to the benchmark model in order to improve the identification. This is

motivated by the coefficient restriction in the linear econometric models. Specifically, we have imposed (a) two Kimball parameters are set to be 10 (they are 'calibrated'), and (b) motivated by Canova and Sala (2009), we force indexation and Calvo parameters are the same for goods price and wage $\iota_p = \iota_w$ and $\xi_p = \xi_w$. These parameter restrictions, of course, reduce the number of free parameters to be estimated. In the similar vein to this exercise, 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.[15]

The normalized posterior variances for the restricted model are given in Table 7 and the Identification Ratios of the restricted model is presented in Table 8. The main findings are as follows. First, not surprisingly, the identification of $\iota_p = \iota_w$ and $\xi_p = \xi_w$ have improved dramatically. Second, the IRs for the other parameters are also improved, although their IRs are still relatively low. Monetary policy parameters r_π and r_y do not improve very much, though. Third, general tendency that we discussed above still holds here again. For example, the trend growth rate is quite strongly identified, and parameters for the exogenous processes are relatively well identified.

4 Further Issues in using the diagnostic indicator

In this section, we briefly discuss some additional issues for the applied researcher to consider in the estimation of structural models. First, in terms of the choice between the Hessian and the posterior variance derived from the MCMC, we suggest that the use of the Hessian is to be preferred. As we have shown, the results are almost identical and as the additional computational burden to obtain the Hessian is much lower than MCMC re-sampling.

Related to this, in this paper, to facilitate comparison, the Hessian is inverted in Table 7, 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 of estimates: i.e., examine the Hessian directly without inverting it. This step may be particularly

[15] We thank Hashem Pesaran for this suggestion, which we commend to future work.

important in the case of weak identification where the Hessian may become near singular (or ill-conditioned). In such a case, we can examine the normalized precision, which is simply the diagonal elements of the Hessian divided by the sample size T.

Second, not surprisingly, if sample size is too few, there is stochastic variation, which reduces the reliability of the KPS indicator. Also, in the spirit of KPS, we are interested in the data identification. Hence, if the sample size is too small, almost all parameters are identified merely because of the prior. We need a large sample size to eliminate the effect of prior.

Third, there must be enough gap between n (shorter data) and N (longer data). In Tables 4 and 6, we show the comparison between $n = 5,000$ and $N = 10,000$, but their results are less similar to the results of $n = 1,000$ and $N = 10,000$. Actually, the results of $n = 1,000$ and $N = 5,000$ are much more closer to those of $n = 1,000$ and $N = 10,000$. We suspect that if n and N are too close, the indicator does not work very well.

Fourth, we ought to be chary in any application of a mechanical cut-off rule for the cardinal value of IR at 1. 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 reasons are, as discussed above, (a) we still have, though very large, finite sample size, and (b) the role of the prior is greater for smaller sample size. In our exercise, we used 0.8 to demarcate weakly and well identified parameters, but this number should be model dependent.

4.1 Sensitivity

We undertake two sensitivity analyses in this section. First we examine our results with an additional random seed and then compare the results of the inverse Hessian to the use of the Hessian directly. Figure 3 compares two identical uses of the KPS indicator for different sample data. In this exercise, we use the same Smets and Wouters' (2007) model to generate two artificial data sets by using two different random number seeds. They are not surprisingly correlated to each other (correlation is 0.68). However, the

finding that two Kimball curvature parameters are not identified is still clear for both random seeds.[16]

Figure 4 compares the inverse Hessian based (variance-covariance based) and the Hessian-based (precision based). As discussed previously, from the computational viewpoint, we recommend to use the inverse Hessian rather than MCMC to obtain asymptotic variance. However, if so, a natural conjecture is, to avoid inversion of a (potentially) big matrix, we could use Hessian directly to investigate asymptotic precision. Again, what we have found is, though there is a significant correlation between the variance based and the precision based methods, there are non trivial discrepancy between them.

We can summarize what we have found as follows. First, we conclude that the only parameters that are not identified are the two Kimball parameters. This finding is quite robust throughout our experiments. Secondly, the Hessian based method is much better than the MCMC method, because, while the results are identical, the computation time is much shorter for the Hessian method. Third, we should not use 1 as a threshold value for IR. Rather, as the IRs of the two Kimball curvature parameters show, if IR is near $n/T = 1,000/10,000$, it could be the evidence that the parameter is unidentified.

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 data identification, i.e. identification without the help or dominance of the Bayesian prior. There may be some use in combining the KPS method 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

[16] To aid comparison we have removed the observations with a very high IR. The underlying tables are available on request.

and *vice versa*. 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 do not typically rely on the data.

In our simple experiments, we find that many parameters in the Smets and Wouters (2007) model, which works as a benchmark in many DSGE applications, are identified but some parameters related to (a) level, (b) monetary policy rule and (c) price and wage stickiness may pose more of a problem. These findings are consistent with those in the emerging literature and are also clearly demonstrated by the KPS measure of posterior precision, as represented by the Identification Ratio. Perhaps the clearest finding is that two Kimball curvature parameters are unidentified, consistent with Iskrev (2010a). Researchers may consider presenting the KPS statistic for every estimation presented of a DSGE model so that better more meaningful inference can be derived as to the usefulness or otherwise of the estimates for further analytical work.

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 $T = 1,000$ or $5,000$ 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 can be shown to be simply identified might be considered a great comfort to those using Bayesian estimation.

References

[1] An, S. and F. Schorfheide (2007). "Bayesian Analysis of DSGE Models." *Econometric reviews*, 26, pp.113-172.

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

[3] Canova, F. and L. Sala (2009). "Back to Square One: Identification Issues in DSGE Models." *Journal of Monetary Economics*, 56, pp.431-449.

[4] Consolo, A., C. A. Favero, and A. Paccagnini (2009). "On the Statistical Identification of DSGE Models." *Journal of Econometrics*, 150, pp.99-115.

[5] Iskrev, N. (2008). "Evaluating the Information Matrix in Linearized DSGE Models." *Economics Letters*, 99, pp.607-610.

[6] Iskrev, N. (2010a). "Local Identification in DSGE Models." *Journal of Monetary Economics*, 57, pp.189-202.

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

[8] Iskrev, N. and M. Ratto (2010). "Analysing Identification Issues in DSGE Models." Mimeo, MONFISPOL papers, Stressa, Italy.

[9] Komunjer, I. and S. Ng (2011). "Dynamic Identification of DSGE Models." *Econometrica*, 79(6), pp.1995–2032.

[10] Koop, G., M. H. Pesaran, and R. P. Smith (2013). "On Identification of Bayesian DSGE Models." *Journal of Business and Economic Statistics*, 31, pp.300-314.

[11] Ratto, M., S. Tarantola, A. Saltelli and P. C. Young (2004). "Accelerated Estimation of Sensitivity Indices Using State Dependent Parameter Models.", In K. M. Hanson

and F. M. Hemez (eds), Sensitivity Analysis of Model Output, Proceedings of the 4th International Conference on Sensitivity Analysis of Model Output.

[12] Ratto, M. (2008). "Analysing DSGE Models with Global Sensitivity Analysis." *Computational Economics*, 31, pp.115-139.

[13] 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*, 26, pp.222-233.

[14] Sims, Christopher A, (2002). "Solving Linear Rational Expectations Models," *Computational Economics*, 20, pp.1-20.

[15] Smets, F. and R. Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, 97, pp.586-606.

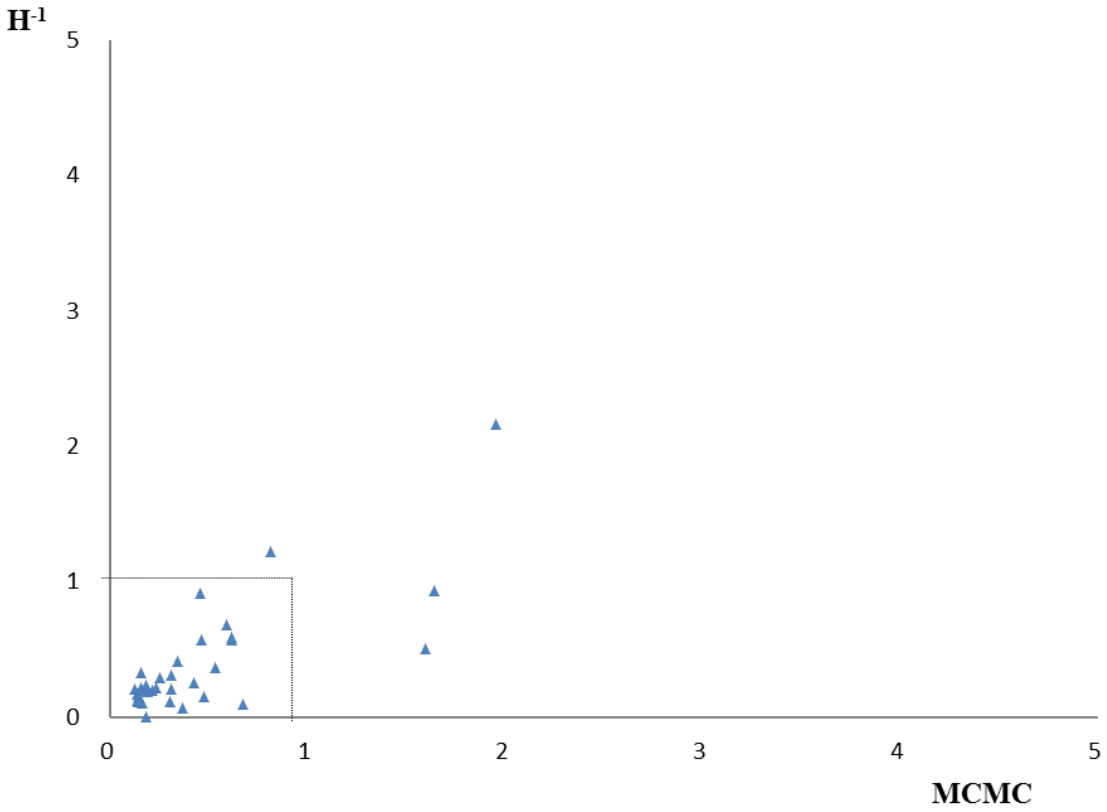

Figure 1: Identification Ratio of MCMC and Inverse HEssian for $\frac{T=10}{T=100}$

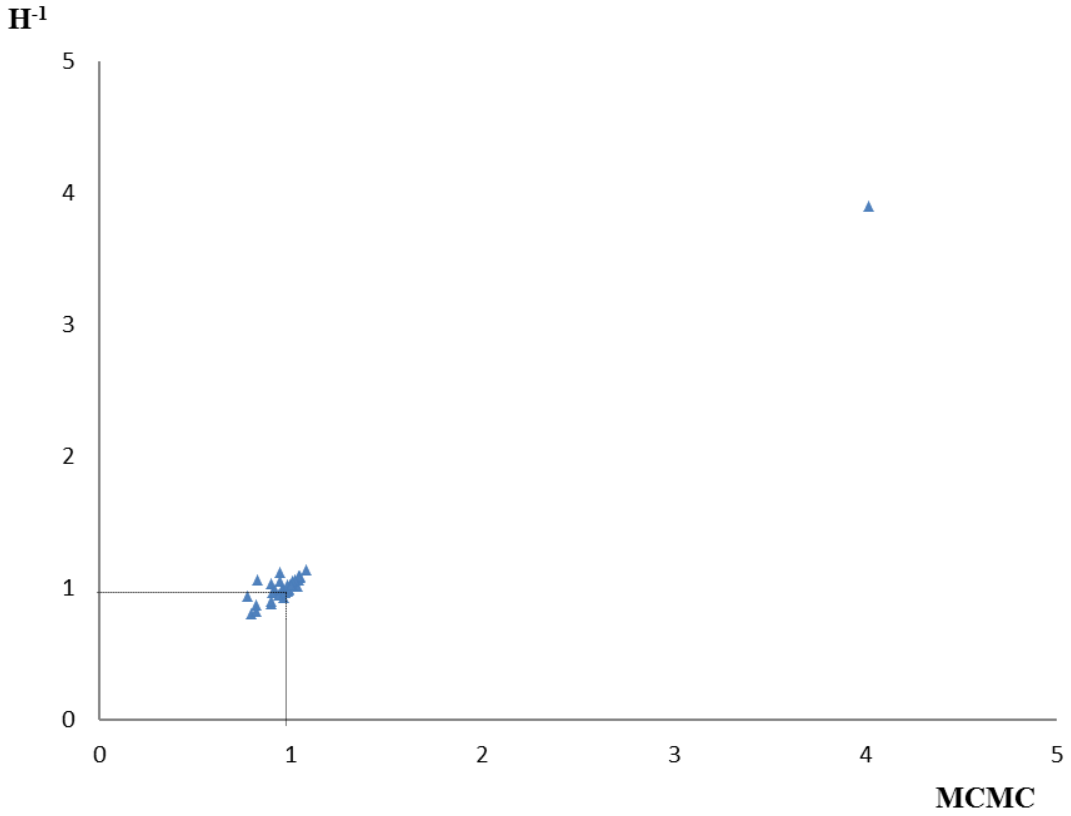

Figure 2: Identification Ratio of MCMC and Inverse HEssian for $\frac{T=5,000}{T=10,000}$

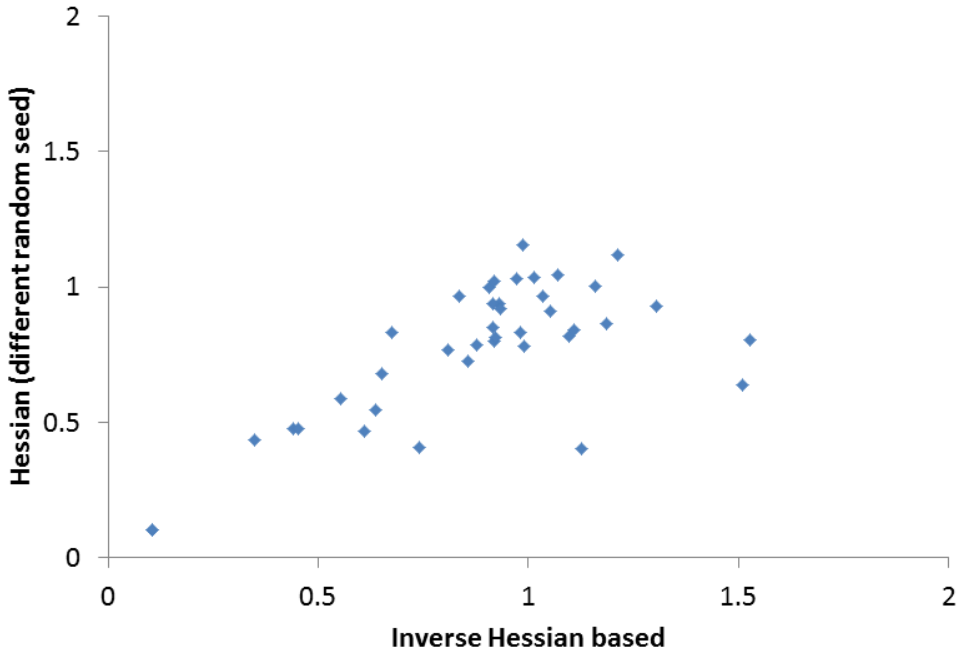

Figure 3: Hessian measure with alternate random seed

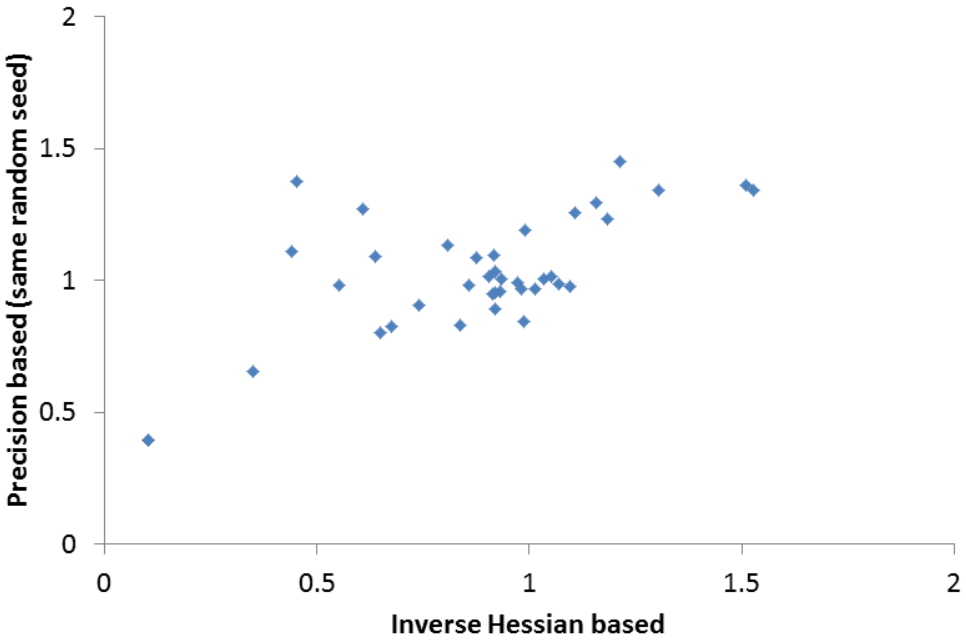

Figure 4: Inverse Hessian compared to precision measure

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 method)

Parameter	$T = 10$	$T = 100$	$T = 1,000$	$T = 5,000$	$T = 10,000$
φ	15.228	123.098	388.638	580.608	705.941
σ_c	0.982	1.600	4.399	7.434	9.576
h	0.079	0.260	0.412	0.521	0.481
ξ_w	0.076	0.416	1.673	2.888	3.209
σ_l	4.958	32.318	108.535	169.368	184.441
ξ_p	0.041	0.272	0.580	0.640	0.804
ι_w	0.223	1.628	5.004	5.903	5.852
ι_p	0.227	0.921	1.391	2.426	2.540
Ψ	0.201	1.504	2.750	2.308	2.560
ϕ	0.153	0.723	2.137	2.398	2.560
r_π	0.527	3.777	13.923	21.490	21.875
ρ_r	0.053	0.115	0.216	0.249	0.264
r_y	0.024	0.152	0.186	0.225	0.225
$r_{\Delta y}$	0.023	0.099	0.275	0.334	0.333
$\overline{\pi}$	0.105	0.674	4.533	9.725	11.902
β	0.076	0.542	1.579	2.821	3.409
$\bar{l}$	42.190	52.170	238.712	404.481	450.140
γ	0.086	0.044	0.000	0.000	0.000
α	0.020	0.060	0.073	0.064	0.066
ρ_a	0.277	0.045	0.018	0.027	0.025
ρ_b	0.280	0.911	1.102	1.214	1.169
ρ_g	0.332	0.050	0.011	0.013	0.013
ρ_{qs}	0.388	0.731	0.632	0.657	0.641
ρ_{ms}	0.365	0.801	1.176	1.181	1.253
ρ_π	0.450	0.735	0.439	0.347	0.387
ρ_w	0.192	1.075	0.036	0.027	0.027
ω_p	0.400	2.272	2.706	2.610	2.710
ω_g	0.627	1.064	1.729	1.564	1.612
ω_w	0.293	2.142	0.287	0.266	0.268
σ_a	0.075	0.159	0.165	0.167	0.174
σ_b	0.026	0.072	0.106	0.114	0.109
σ_g	0.267	0.163	0.183	0.168	0.169
σ_{qs}	0.157	0.520	0.357	0.379	0.365
σ_{ms}	0.010	0.024	0.043	0.042	0.041
σ_π	0.057	0.036	0.062	0.067	0.068
σ_w	0.043	0.064	0.081	0.089	0.089

Note: 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: Identification Ratio (structural parameter generated by MCMC method)

$\frac{T=10}{T=100}$		$\frac{T=100}{T=1,000}$		$\frac{T=1,000}{T=5,000}$		$\frac{T=1,000}{T=10,000}$		$\frac{T=5,000}{T=10,000}$	
ρ_g	6.636	γ	104.334	γ	32.334	γ	129.940	γ	4.019
ρ_a	6.220	ρ_w	29.847	ρ_w	1.340	ρ_w	1.318	h	1.083
γ	1.950	ω_w	7.461	ρ_π	1.267	ρ_π	1.134	ρ_a	1.047
σ_g	1.640	ρ_g	4.669	Ψ	1.192	α	1.110	σ_b	1.045
σ_π	1.595	ρ_a	2.474	α	1.147	σ_g	1.081	σ_{ms}	1.042
$\bar{l}$	0.809	ρ_π	1.674	ω_g	1.105	Ψ	1.074	ρ_b	1.039
σ_w	0.669	σ_{qs}	1.457	σ_g	1.091	ω_g	1.072	σ_{qs}	1.037
σ_c	0.614	ρ_{qs}	1.156	ω_w	1.082	ω_w	1.072	ρ_{qs}	1.025
ρ_π	0.612	σ_a	0.963	ω_p	1.037	σ_{ms}	1.053	ρ_g	1.012
ω_g	0.589	σ_g	0.891	σ_{ms}	1.011	ω_p	0.999	ι_w	1.009
ρ_{qs}	0.531	ω_p	0.840	ρ_{ms}	0.996	ρ_{qs}	0.986	$r_{\Delta y}$	1.005
σ_a	0.474	ρ_b	0.827	σ_a	0.990	σ_{qs}	0.978	r_y	1.001
ρ_r	0.459	α	0.823	ρ_{qs}	0.961	σ_b	0.973	σ_w	0.998
ρ_{ms}	0.456	r_y	0.817	σ_{qs}	0.943	σ_a	0.949	ω_w	0.991
σ_{ms}	0.419	σ_w	0.786	σ_π	0.935	ρ_b	0.943	σ_g	0.991
σ_b	0.361	σ_b	0.682	σ_b	0.931	ρ_{ms}	0.939	ρ_w	0.984
α	0.340	ρ_{ms}	0.681	σ_w	0.908	σ_π	0.910	r_π	0.982
ρ_b	0.308	ι_p	0.662	ρ_b	0.908	σ_w	0.906	σ_π	0.973
h	0.306	h	0.631	ξ_p	0.907	h	0.856	ω_g	0.970
σ_{qs}	0.302	ω_g	0.616	ϕ	0.891	ι_w	0.855	α	0.968
ι_p	0.246	σ_π	0.575	ρ_r	0.866	ϕ	0.835	ω_p	0.963
$r_{\Delta y}$	0.232	σ_{ms}	0.554	ι_w	0.848	r_y	0.828	σ_a	0.959
ϕ	0.211	Ψ	0.547	r_y	0.827	$r_{\Delta y}$	0.827	ι_p	0.955
ξ_w	0.182	ρ_r	0.535	$r_{\Delta y}$	0.823	ρ_g	0.825	ρ_{ms}	0.942
ρ_w	0.178	ξ_p	0.469	ρ_g	0.815	ρ_r	0.816	ρ_r	0.942
ω_p	0.176	σ_c	0.364	h	0.790	ξ_p	0.722	ϕ	0.937
r_y	0.157	$r_{\Delta y}$	0.361	ρ_a	0.680	ρ_a	0.712	σ_l	0.918
$\overline{\pi}$	0.155	β	0.343	φ	0.669	r_π	0.637	Ψ	0.902
σ_l	0.153	ϕ	0.339	r_π	0.648	σ_l	0.589	ξ_w	0.900
ξ_p	0.151	ι_w	0.325	σ_l	0.641	φ	0.551	$\bar{l}$	0.899
β	0.140	φ	0.317	σ_c	0.592	ι_p	0.548	ρ_π	0.896
r_π	0.140	σ_l	0.298	$\bar{l}$	0.590	$\bar{l}$	0.530	β	0.828
ι_w	0.137	r_π	0.271	ξ_w	0.579	ξ_w	0.521	φ	0.823
ω_w	0.137	ξ_w	0.249	ι_p	0.573	β	0.463	$\overline{\pi}$	0.817
Ψ	0.133	$\bar{l}$	0.219	β	0.560	σ_c	0.459	ξ_p	0.796
φ	0.124	$\overline{\pi}$	0.149	$\overline{\pi}$	0.466	$\overline{\pi}$	0.381	σ_c	0.776

Note: 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 5: Normalized posterior variances of structural parameters (generated by H^{-1} method)

Parameter	$T = 10$	$T = 100$	$T = 1,000$	$T = 5,000$	$T = 10,000$
φ	27.608	132.429	355.752	565.491	649.283
σ_c	0.963	1.639	4.735	7.712	8.214
h	0.054	0.266	0.431	0.527	0.464
ξ_w	0.079	0.426	1.730	3.001	2.903
σ_l	7.648	34.600	103.932	159.983	162.949
ξ_p	0.096	0.288	0.565	0.635	0.789
ι_w	0.330	1.914	5.100	5.898	5.608
ι_p	0.283	0.965	1.501	2.414	2.442
Ψ	0.213	1.752	2.875	2.351	2.443
ϕ	0.150	0.758	2.081	2.336	2.459
r_π	0.565	3.689	13.964	21.550	21.082
ρ_r	0.065	0.114	0.225	0.256	0.244
r_y	0.016	0.152	0.187	0.227	0.221
$r_{\Delta y}$	0.022	0.101	0.270	0.330	0.322
$\overline{\pi}$	0.114	0.565	4.606	9.590	11.700
β	0.066	0.555	1.630	2.766	2.609
$\bar{l}$	46.177	37.808	239.787	402.043	449.101
γ	0.093	0.043	0.000	0.000	0.000
α	0.024	0.059	0.070	0.061	0.062
ρ_a	0.320	0.029	0.017	0.026	0.024
ρ_b	0.284	0.928	1.111	1.175	1.081
ρ_g	0.488	0.047	0.010	0.013	0.013
ρ_{qs}	0.289	0.791	0.665	0.691	0.650
ρ_{ms}	0.751	0.823	1.258	1.257	1.124
ρ_π	0.390	0.683	0.402	0.338	0.383
ρ_w	0.005	1.115	0.033	0.026	0.026
ω_p	0.734	3.064	2.485	2.427	2.609
ω_g	0.679	0.994	1.738	1.556	1.610
ω_w	0.426	2.559	0.275	0.263	0.264
σ_a	0.025	0.162	0.158	0.161	0.162
σ_b	0.005	0.069	0.106	0.111	0.102
σ_g	0.142	0.152	0.179	0.165	0.169
σ_{qs}	0.051	0.455	0.350	0.373	0.368
σ_{ms}	0.006	0.022	0.043	0.044	0.041
σ_π	0.019	0.038	0.062	0.066	0.068
σ_w	0.006	0.063	0.082	0.090	0.088

Note: 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 6: Identification Ratio (structural parameter generated by H^{-1} method)

$\frac{T=10}{T=100}$		$\frac{T=100}{T=1,000}$		$\frac{T=1,000}{T=5,000}$		$\frac{T=1,000}{T=10,000}$		$\frac{T=5,000}{T=10,000}$	
γ	11.108	γ	105.824	γ	32.153	γ	125.488	γ	3.903
ρ_g	10.492	ρ_w	34.115	ρ_w	1.276	ρ_w	1.241	h	1.134
γ	2.160	ω_w	9.306	Ψ	1.223	Ψ	1.177	ρ_{ms}	1.119
$\bar{l}$	1.221	ρ_g	4.477	ρ_π	1.190	α	1.120	ρ_b	1.087
σ_g	0.933	ρ_a	1.713	α	1.154	ρ_{ms}	1.119	σ_b	1.086
ρ_{ms}	0.912	ρ_π	1.700	ω_g	1.117	ω_g	1.080	ρ_a	1.077
ω_g	0.683	σ_{qs}	1.302	σ_g	1.082	σ_g	1.061	ρ_{qs}	1.063
σ_c	0.588	ω_p	1.233	ω_w	1.045	ρ_π	1.048	σ_{ms}	1.061
ρ_r	0.572	ρ_{qs}	1.189	ω_p	1.024	σ_{ms}	1.045	β	1.060
ρ_π	0.571	σ_a	1.021	ρ_{ms}	1.000	ω_w	1.040	ι_w	1.052
σ_π	0.504	σ_g	0.851	σ_{ms}	0.985	σ_b	1.038	ρ_r	1.050
α	0.412	α	0.848	σ_a	0.980	ρ_b	1.028	ξ_w	1.034
ρ_{qs}	0.365	ρ_b	0.835	ρ_{qs}	0.963	ρ_{qs}	1.024	$r_{\Delta y}$	1.028
ξ_p	0.332	r_y	0.816	σ_b	0.956	σ_a	0.975	r_y	1.027
ρ_b	0.306	σ_w	0.770	ρ_b	0.946	ω_p	0.952	r_π	1.022
ι_p	0.293	ρ_{ms}	0.655	σ_π	0.941	σ_{qs}	0.951	σ_w	1.022
σ_{ms}	0.255	σ_b	0.653	σ_{qs}	0.938	σ_w	0.929	ρ_g	1.020
ω_p	0.239	ι_p	0.643	σ_w	0.909	h	0.928	σ_{qs}	1.013
σ_l	0.221	h	0.616	ϕ	0.891	ρ_r	0.923	ω_w	0.995
$r_{\Delta y}$	0.214	σ_π	0.611	ξ_p	0.891	σ_π	0.918	σ_a	0.995
φ	0.208	Ψ	0.609	ρ_r	0.879	ι_w	0.909	ι_p	0.989
h	0.205	ω_g	0.572	ι_w	0.865	ϕ	0.846	σ_l	0.982
$\overline{\pi}$	0.202	ξ_p	0.510	r_y	0.823	r_y	0.846	σ_g	0.981
ϕ	0.198	σ_{ms}	0.509	h	0.819	$r_{\Delta y}$	0.841	σ_π	0.976
ξ_w	0.186	ρ_r	0.504	$r_{\Delta y}$	0.818	ρ_g	0.806	ρ_w	0.973
ι_w	0.172	ι_w	0.375	ρ_g	0.790	ξ_p	0.717	α	0.970
ω_w	0.166	$r_{\Delta y}$	0.372	ρ_a	0.650	ρ_a	0.700	ω_g	0.967
σ_a	0.155	φ	0.372	σ_l	0.650	r_π	0.662	Ψ	0.963
r_π	0.153	ϕ	0.364	r_π	0.648	σ_l	0.638	ϕ	0.950
Ψ	0.122	σ_c	0.346	φ	0.629	β	0.625	σ_c	0.939
β	0.118	β	0.341	ι_p	0.622	ι_p	0.615	ω_p	0.930
σ_{qs}	0.113	σ_l	0.333	σ_c	0.614	ξ_w	0.596	$\bar{l}$	0.895
r_y	0.104	r_π	0.264	$\bar{l}$	0.596	σ_c	0.577	ρ_π	0.881
σ_w	0.097	ξ_w	0.246	β	0.589	φ	0.548	φ	0.871
σ_b	0.071	$\bar{l}$	0.158	ξ_w	0.577	$\bar{l}$	0.534	$\overline{\pi}$	0.820
ρ_w	0.005	$\overline{\pi}$	0.123	$\overline{\pi}$	0.480	$\overline{\pi}$	0.394	ξ_p	0.805

Note: 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 7: Normalized posterior variances of the restricted model

Parameter	$T = 10$	$T = 100$	$T = 1,000$	$T = 5,000$	$T = 10,000$
σ_c	0.655	2.373	3.825	6.883	5.634
h	0.088	0.372	0.335	0.407	0.425
σ_l	4.717	37.022	124.942	221.464	194.384
ξ_p	0.034	0.364	0.349	0.390	0.382
ι_p	0.045	0.876	1.584	2.029	1.946
Ψ	0.229	1.062	2.224	1.891	1.820
ϕ	0.063	0.686	1.805	2.442	2.452
r_π	0.236	2.649	9.502	17.370	19.702
γ	0.089	0.039	0.000	0.000	0.000
α	0.011	0.023	0.063	0.059	0.059
ρ_a	0.016	0.020	0.019	0.022	0.022
ρ_b	0.308	1.175	0.993	1.146	1.106
ρ_g	0.127	0.303	0.017	0.015	0.015
ρ_{qs}	0.233	0.963	0.756	0.651	0.652
ρ_{ms}	0.228	0.617	1.104	1.285	1.070
ρ_π	0.029	2.253	0.777	0.992	0.948
ρ_w	0.074	0.228	0.016	0.020	0.022
ω_p	0.327	4.027	2.057	2.923	2.625
ω_g	0.080	0.541	1.097	1.309	1.298
ω_w	0.215	1.088	0.213	0.165	0.177
σ_a	0.422	0.149	0.177	0.172	0.178
σ_b	0.020	0.079	0.101	0.112	0.107
σ_g	0.025	0.075	0.138	0.157	0.160
σ_{qs}	0.202	0.420	0.474	0.399	0.387
σ_{ms}	0.003	0.024	0.032	0.036	0.038
σ_π	0.001	0.025	0.037	0.044	0.041
σ_w	0.005	0.070	0.059	0.060	0.060

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 $\xi_w = \xi_p$ and $\iota_w = \iota_p$.

Table 8: Identification Ratio (the restricted model)

$\frac{T=10}{T=100}$		$\frac{T=100}{T=1,000}$		$\frac{T=1,000}{T=5,000}$		$\frac{T=1,000}{T=10,000}$		$\frac{T=5,000}{T=10,000}$	
σ_a	2.825	γ	91.604	γ	35.955	γ	135.608	γ	3.772
γ	2.295	ρ_g	17.435	ω_w	1.293	σ_{qs}	1.225	σ_c	1.222
ρ_a	0.779	ρ_w	14.398	σ_{qs}	1.190	Ψ	1.222	ρ_{ms}	1.201
σ_{qs}	0.480	ω_w	5.097	Ψ	1.176	ω_w	1.206	σ_l	1.139
α	0.474	ρ_π	2.901	ρ_{qs}	1.162	ρ_{qs}	1.160	ω_p	1.114
ρ_g	0.421	ω_p	1.958	ρ_g	1.135	ρ_g	1.129	σ_π	1.054
ρ_{ms}	0.370	ρ_{qs}	1.274	α	1.077	α	1.074	σ_b	1.051
σ_g	0.328	ρ_b	1.184	σ_a	1.030	ρ_{ms}	1.032	ρ_π	1.047
ρ_w	0.326	σ_w	1.180	σ_w	0.988	σ_a	0.994	ι_p	1.043
σ_c	0.276	h	1.110	σ_b	0.901	σ_w	0.986	Ψ	1.039
ρ_b	0.262	ρ_a	1.074	ξ_p	0.897	σ_b	0.947	ρ_b	1.036
σ_b	0.255	ξ_p	1.043	σ_{ms}	0.894	ξ_p	0.914	σ_{qs}	1.030
ρ_{qs}	0.242	σ_{qs}	0.885	σ_g	0.880	σ_π	0.900	ξ_p	1.019
h	0.237	σ_a	0.842	ρ_b	0.866	ρ_b	0.897	ω_g	1.008
Ψ	0.216	σ_b	0.781	ρ_{ms}	0.859	σ_{ms}	0.862	ρ_a	1.000
ω_w	0.197	σ_{ms}	0.725	σ_π	0.854	σ_g	0.861	ρ_{qs}	0.999
ω_g	0.147	σ_π	0.684	ω_g	0.838	ω_g	0.845	α	0.997
σ_l	0.127	σ_c	0.620	ρ_a	0.827	ρ_a	0.827	σ_w	0.997
σ_{ms}	0.119	ρ_{ms}	0.559	h	0.822	ρ_π	0.820	ϕ	0.996
ϕ	0.092	ι_p	0.553	ρ_π	0.783	ι_p	0.814	ρ_g	0.994
ξ_p	0.092	σ_g	0.546	ι_p	0.781	h	0.788	σ_g	0.979
r_π	0.089	ω_g	0.494	ρ_w	0.776	ω_p	0.784	σ_a	0.965
ω_p	0.081	Ψ	0.478	ϕ	0.739	ϕ	0.736	σ_{ms}	0.964
σ_w	0.073	ϕ	0.380	ω_p	0.704	ρ_w	0.718	h	0.959
ι_p	0.051	α	0.367	σ_l	0.564	σ_c	0.679	ω_w	0.933
σ_π	0.028	σ_l	0.296	σ_c	0.556	σ_l	0.643	ρ_w	0.925
ρ_π	0.013	r_π	0.279	r_π	0.547	r_π	0.482	r_π	0.882

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 $\overline{l}$, 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 $\xi_w = \xi_p$ and $\iota_w = \iota_p$.