How much do we learn from the estimation of DSGE models? A case study of identification issues in a New Keynesian business cycle model*

Nikolay Iskrev
Department of Economics, University of Michigan
niskrev@umich.edu

October 1, 2007

`preliminary and incomplete`
`comments solicited`

Abstract

This paper proposes a new approach for studying parameter identification in linearized DSGE models, based on analytical evaluation of the Information matrix of such models. The Information matrix is decomposed into a part that depends on the model only, and a part which also depends on the data used for estimation. This allows researchers to determine: first, whether the parameters of the model are identified; second, whether identification is strong or weak; and third, if identification problems are detected, whether they originate in the structure of the model, or in the data. We apply this approach to study parameter identification in a large-scale monetary business cycle model estimated by Smets and Wouters (2007). We find that, for parameters that are identifiable, identification is generally very weak. Moreover our results indicate that the problem is largely embedded in the structure of the model, and, therefore, cannot be resolved by using more informative data. We also show that there are substantial differences in the parameters estimates obtained with classical and Bayesian estimation methods. We conclude that using estimated DSGE models for policy analysis should be done with caution since, when identification is weak, the results are likely to be strongly influenced by the prior distribution.

Keywords: DSGE models, Identification, Information matrix

JEL classification: C32, C51, C52, E32

---

*This paper is a substantial revision of "Parameter Identification in the Smets and Wouters model", and a paper with the same title presented at the Third Dynare conference, Paris, September 2007. The previous version is available at `http://www-personal.umich.edu/∼niskrev/`

Contents

... A lot of your posteriors look exactly like the priors...

Richard Blundell, when awarding Frank Smets and Raf Wouters with the Hicks-Tinbergen Medal at the 2004 EEA Meetings.

1 Introduction

The last several years have witnessed a remarkable growth in the research on empirical evaluation of DSGE models. Nowadays researchers routinely estimate rich micro-founded models, that until recently had to be calibrated. Unlike reduced-form or single equation estimation methods, the full set of model parameters are being estimated in an internally-consistent fashion. This, together with the finding that empirical DSGE models can fit the data as well as model-free reduced-form vector autoregressions (VAR), has made them extremely popular in central banks and other policy-making institutions.[1]

A question which is rarely addressed in the empirical DSGE literature is that of parameter identifiability. This is surprising as identification is a prerequisite for estimation of the parameters, and the ability to do that for full-fledged structural models is believed to be one of the main accomplishments of this line of research. That parameter identification is a potentially serious issue for DSGE models is not a new concern. Among the authors who have made this point are Sargent (1976) and Pesaran (1989). More recently Beyer and Farmer (2004) provide several examples of commonly used models that are unidentifiable. They argue that the problem is likely to be common in DSGE models.

In most empirical DSGE papers the question of parameter identification is not confronted directly. Usually, if some of the parameters are considered to be of lesser interest, and/or with potentially problematic identifiability, their values are calibrated and assumed known, instead of being estimated. Furthermore, since DSGE models are frequently estimated using Bayesian methods, potential identification problems remain hidden due to the use of priors. As a result, it is often unclear to what extent the reported estimates reflect information in the data instead of subjective beliefs or other considerations reflected in the choice of prior distribution for the parameters. One reason why this is an important issue is that DSGE models are increasingly being used for analyzing policy-relevant questions, such as, for instance, the design of optimal monetary policy. Such analysis often hinges crucially on the values assigned to the parameters of the model. It is, therefore, important to know how informative is the data for the parameters of interest, and whether there are any benefits from estimating instead of calibrating the models we use to address policy questions.

The objective of this paper is to shed light on the relative importance of information from the data versus subjective prior beliefs for the estimation of a state-of-the-art DSGE model. We address the problem in two steps. First, we develop a new identification analysis procedure, based on analytical evaluation of the information matrix, and use it to study the identification of the parameters in the model. Second, we estimate the model using maximum likelihood, and compare the results to those obtained by using Bayesian techniques. The model we consider is a large-scale New Keynesian business cycle model with various real and nominal frictions developed and estimated in Smets and Wouters (2007). Similar models have been studied, using Bayesian techniques, in Onatski and Williams (2004), DelNegro, Schorfheide, Smets, and Wouters (2005),

[1]Models similar to the one considered in this paper have been estimated, and are used for policy analysis in institutions such as the Federal Reserve Board, the European Central Bank, Bank of England, RiksBank, the Bank of Canada, and the IMF.

Justiniano and Primiceri (2006), and Boivin and Giannoni (2006). Previous research suggests that the model fits the data well, in some cases outperforming unrestricted vector autoregressions in out-of-sample forecasting. Schmitt-Grohe and Uribe (2004) and Levin, Onatski, Williams, and Williams (2005) study the design of optimal monetary policy rules in estimated versions of that model (see also Juillard, Karam, Laxton, and Pesenti (2006)).

Although its importance has been recognized (see e.g. An and Schorfheide (2005)), the identification of parameters in this or other similar DSGE models has not been studied previously. Perhaps the main reason for this is that applying the standard approach to identification is very difficult for DSGE models (see Schorfheide (2007)). In general, DSGE models have the following form:

$$E_t J(\tilde{Z}_{t+1}, \tilde{Z}_t, \tilde{Z}_{t-1}, U_t; \theta) = 0 \tag{1.1}$$

where J is a non-linear function of the endogenous variables $\tilde{Z}$, and the exogenous shocks U, and θ is a vector of deep parameter. Since the model in (1.1) is, for most purposes, too difficult to work with, researchers use a linear or log-linear approximation of (1.1) around the steady state. The resulting system of linear stochastic equations is of the form

$$E_t \hat{J}(Z_{t+1}, Z_t, Z_{t-1}, U_t; \theta) = 0 \tag{1.2}$$

where Z is the log-deviation of $\tilde{Z}$ from its steady-state level, and $\hat{J}$ is function linear in the variables Z and U. Solving the linearized version of the DSGE model yields a reduced form model, given by

$$R(Z_t, Z_{t-1}, U_t; \tau) = 0 \tag{1.3}$$

where R is a linear function of Z and U, parameterized by the vector of reduced form parameters τ.

A classic result of Rothenberg (1971) relates the identification of parameters to the information matrix of the model. In particular, a singular information matrix indicates that some parameters in θ are not identifiable. Finding the information matrix in DSGE models, however, is not straightforward since, for most models, the mapping from the structural model - (1.2) to the reduced-form one - (1.3), can only be found numerically. This makes the analytical derivation of the information matrix by direct differentiation of the likelihood function impossible.

In Iskrev (2007a) we showed how the information matrix can be evaluated analytically for linearized DSGE models. We factorize the information matrix for θ as a product of two terms: one is the gradient of the mapping from reduced from parameters τ to deep parameters θ; the second is the information matrix of the reduced-form model (1.3). Both factors can be derived and evaluated analytically. This approach not only makes a precise evaluation of the information matrix possible, but also provides a necessary condition for identification of the deep parameters, which does not depend on the data. The condition is that the the gradient of the mapping from θ to τ has a full rank. This mapping is completely independent from the data used in estimation. Thus, we can detect identification problems that are inherent in the structure of the DSGE model, and not caused by data deficiencies. In addition, we can determine which parameters cause the identification problems, and, possibly, find a better identified parametrization of the model we are interested in estimating.

Following this approach, we conduct a thorough identification analysis of the model we consider in our case study. In particular, we draw a large number of points from the parameter space of the model, and check the necessary and sufficient rank conditions at each one of them. In addition, we evaluate the conditioning of the matrices whose ranks determine identification. In doing so we establish not only whether the parameters are identifiable in the strict sense, but also how strong identification is. We do this for six parameterizations that differ in which parameters are assumed to be known. Then we turn to the estimation of parameters using quarterly US data. We depart from the previous studies in using maximum likelihood for estimation of the model. This allows us to compare parameter estimates driven by the data only, with those obtained with Bayesian methods, which are determined by both the data and the prior distribution. When the number of observations is large, the two approaches should produce similar results. In small samples, however, the prior distribution could be very influential, especially when identification is weak. This may result in parameter estimates that have little to do with the actual data used for estimation.

On the identification side, we find that Smets and Wouters (2007), who state that three of the deep parameters of the model are not identifiable, are correct only with respect to two of them. The third one - wage markup parameter, is, in fact, identifiable, although generally very weakly so. When we restrict our analysis to identifiable parameterizations, we find that identification is generally quite weak. We show that this problem is to a large degree embedded in the structure of the model, and thus cannot be resolved by using more informative data. Furthermore, we are able to determine which of the deep parameters are most responsible for the weak identifiability of the model as a whole. We find, for instance, that it is difficult to distinguish between the elasticity of labor supply, the Calvo parameter for wages, and two wage mark up parameters. These four parameters are nearly confounded in most of the parameter space. Similar problem is found with respect to monetary policy rule parameters - it is difficult to distinguish among the response coefficients to inflation, output, and past interest rate.

On the estimation side, we find that disposing with the strong priors used in previous studies affects substantially the estimates of the parameters in the model. This has important implications for the behavior of the model, as we show using impulse response and variance decomposition analysis.

Our paper is not the first to systematically study parameter identification in DSGE models. An important recent contribution that deals exclusively with these issues is Canova and Sala (2006). There are three important differences between their study and the present paper. First, they approach parameter identification from the perspective of a particular limited information estimation method, namely, impulse response matching (see Rotemberg and Woodford (1998), and Altig, Christiano, Eichenbaum, and Linde (2005) for explanation and illustration of this estimation approach). As they recognize, identification failures of that or other limited information methods does not imply that the problems are generic to all estimation methods. In contrast, if identification fails or is weak when a full information approach is used, as we do here, it will remain a problem for any alternative estimation method. Second, unlike this paper which evaluates the information matrix analytically, Canova and Sala (2006) use numerical approximation of the Hessian. It is well known that numerical differentiation could be very imprecise for highly non-linear functions, as is the case with DSGE models.[2] Moreover, with our approach for computing the information matrix, we are able to distinguish between the model structure and the data, as sources of identification problems. Finally, unlike Canova and Sala (2006) who study identification only in the neighborhood of a particular point in the parameter space, we do that for a large number of points drawn randomly from everywhere in the space. Thus we are able to characterize parameter identification as a global instead of a local problem of the theoretical model.

[2]Hansen, McGrattan, and Sargent (1994) also argue in favor of using analytical derivatives when estimating DSGE models

Regarding the effect of priors for Bayesian estimation of DSGE models, results similar to ours are reported in Onatski and Williams (2004). They estimate a similar large-scale New Keynesian model, using European data, and find that greater prior uncertainty results in substantially different parameter estimates, compared to those obtained with the tighter priors common in the empirical DSGE literature. They do not address formally the issue of parameter identifiability, as we do in this paper.

The rest of the paper is organized as follows. The next section provides an overview of the model we study. Section 3 describes our identification approach, and discusses the results from the analysis of parameter identifiability in the model. In section 4 we estimate the model using maximum likelihood, and compare the results to those reported in Smets and Wouters (2007). Section 5 concludes and gives directions for future work.

2 The model

The model in Smets and Wouters (2007) (see also Christiano, Eichenbaum, and Evans (2005)) is an extension of the standard RBC model featuring a number of nominal frictions, such as price and wage stickiness, and real rigidities - habit formation in consumption, investment adjustment cost, monopolistic competition, and variable cost of adjusting capital utilization. In addition, it contains a large number of serially correlated structural shocks. In this section we present a brief outline of the main components of the model. For details see the appendix accompanying Smets and Wouters (2007).

2.1 Households

There is a continuum of households indexed by h, each having the following utility function

$$E_t\left[\sum_{s=0}^{\infty}\beta^s\frac{1}{1-\sigma_C}\left((C_{t+s}(j)-\lambda C_{t+s-1}(j))^{1-\sigma_C}\right)\exp\left(\frac{\sigma_c-1}{1+\sigma_l}L_{t+s}(j)^{1+\sigma_l}\right)\right] \tag{2.1}$$

where $C_{t+s}(j)$ is consumption, $L_{t+s}(j)$ is hours worked; λ is an external habit persistence parameter.

Each household supplies differentiated labor services monopolistically to a continuum of labor markets charging nominal wage denoted with $W_t(j)$; W_t is an index of the nominal wage in the economy.

Households supply homogeneous labor to labor unions (indexed by l), who then sell it to labor packers. Labor services are differentiated by a union, who therefore have market power. Wage setting by unions is subject to nominal rigidities a la Calvo - each period a union can set the nominal wage to the optimal level with constant probability equal to $1-\xi_w$. Unions which cannot adjust their nominal wage optimally, change it according to the following indexation rule

$$W_{t+s}(l)=\gamma W_{t-1}(l)\pi_{t-1}^{\iota_w}\pi_*^{(1-\iota_w)} \tag{2.2}$$

where γ is the deterministic growth rate, ι_w measures the degree of wage indexation to past inflation, and π_* is the steady state rate of inflation.

Labor packers buy differentiated labor services $L_t(l)$ from unions, package and sell composite labor L_t, defined implicitely by

$$\int_0^1 \mathcal{H}\left(\frac{L_t(l)}{L_t};\lambda_{w,t}\right)dl = 1 \tag{2.3}$$

to the intermediate good sector firms. The function $\mathcal{H}$ is increasing, concave, and satisfies $\mathcal{H}(1) = 1$; $\lambda_{w,t}$ is a stochastic exogenous process changing the elasticity of demand, and the wage markup over the marginal disutility from work.

In addition to supplying labor at wage W_t, households rent capital to the firms producing intermediate goods, and earn rent at rate $R_t^K(j)$. Households accumulate physical capital according to the following law of motion:

$$\bar{K}_t(j) = (1-\delta)\bar{K}_{t-1}(j) + \varepsilon_t^i\left[1 - \mathcal{S}\left(\frac{I_t(j)}{I_{t-1}(j)}\right)\right]I_t(j) \tag{2.4}$$

where δ is the rate of depreciation, I_t is gross investment, and the investment adjustment cost function $\mathcal{S}$ satisfies $\mathcal{S}' > 0,\ \ \mathcal{S}'' > 0$, and in steady state $\mathcal{S} = 0,\ \ \mathcal{S}' = 0$. ε_t^I represents the current state of technology for producing capital, and is interpreted as investment-specific technological progress (Greenwood, Hercowitz, and Krusell (2000)).

Households control the utilization rate $Z_t(j)$ of the physical capital they own, and pay $P_t a(Z_t(j))\bar{K}_{t-1}(j)$ in terms of consumption good when the capital intensity is $Z_t(j)$. The income from renting capital to firms is $R_t^k K_t(j)$, where $K_t(j) = Z_t(j)\bar{K}_{t-1}(j)$ is the flow of capital services provided by the existing stock of physical capital $\bar{K}_{t-1}(j)$. The utility function (2.1) is maximized with respect to consumption, hours, investment, and capital utilization, subject to the capital accumulation equation (2.4), and the following the per-period budget constraint

$$\begin{aligned} C_{t+s}(j) + I_{t+s}(j) + \frac{B_{t+s}(j)}{\varepsilon_{t+s}^b R_{t+s} P_{t+s}} - T_{t+s} &= \frac{W_{t+s}^h(j)}{P_{t+s}} L_{t+s}(j) \\ + \left(\frac{R_{t+s}^k Z_{t+s}(j)}{P_{t+s}} - a(Z_{t+s}(j))\right)\bar{K}_{t+s-1}(j) &+ \frac{B_{t+s-1}(j)}{P_{t+s}} + \frac{\Pi_{t+s}(j)}{P_{t+s}} \end{aligned} \tag{2.5}$$

where B_{t+s} is a one-period nominal bond expressed on a discount basis. ε_t^b is an exogenous premium on the bond return, T_{t+s} is lump-sum taxes or subsidies, and Π_{t+s} is profit distributed by the labor union.

2.2 Firms

A perfectly competitive sector produces a single final good used for consumption and investment. The final good is produced from intermediate inputs $Y_t(i)$ using the following technology:

$$\int_0^1 \mathcal{G}\left(\frac{Y_t(i)}{Y_t};\lambda_{p,t}\right)di = 1 \tag{2.6}$$

where $\mathcal{G}$ is increasing, concave, and $\mathcal{G}(1) = 1$; $\lambda_{p,t}$ is exogenous stochastic process affecting the elasticity of substitution between different intermediate goods, also corresponding to markup over marginal cost for intermediate good firms.

Firms maximize profit given by

$$P_t Y_t - \int_0^1 P_t(i) Y_t(i) di \tag{2.7}$$

where $P_t(i)$ is the prize of intermediate good $Y_t(i)$.

Intermediate goods are produced in a monopolistically competitive sector. Each variety i is produced by a single firm using the following technology:

$$Y_t(i) = \varepsilon_t^a K_t(i)^\alpha (\gamma^t L_t(i))^{1-\alpha} - \Phi \gamma^t \tag{2.8}$$

where Φ is fixed cost, ε_t^a denotes total factor productivity, and γ is deterministic growth rate of labor augmenting technology.

As with wages, every period only a fraction $1 - \xi_P$ of intermediate firms can set optimally the price of the good they produce. The remaining ξ_p firms index their prices to past inflation according to:

$$P_t(t) = \gamma P_{t-1}(i) \pi_{t-1}^{\iota_p} \pi_*^{(1-\iota_p)} \tag{2.9}$$

where ι_p measures the degree of price indexation to past inflation.

2.3 The Government

The budget constraint of the government is

$$P_t G_t + B_{t-1} = T_t + \frac{B_t}{R_t} \tag{2.10}$$

where G_t is government consumption in terms of final good.

The central bank sets the nominal interest rate according to the following rule

$$\frac{R_t}{R^*} = r_t \left(\frac{R_{t-1}}{R^*}\right)^\rho \left[\left(\frac{\pi_t}{\pi^*}\right)^{r_\pi} \left(\frac{Y_t}{Y_t^*}\right)^{r_y}\right]^{1-\rho} \left(\frac{Y_t / Y_{t-1}}{Y_t^* / Y_{t-1}^*}\right)^{r_{\triangle y}} \tag{2.11}$$

where R^* is the steady state level of the gross nominal interest rate, and r_t is a monetary policy shock; Y^* is the potential level of output, defined as the output in a flexible price and wage economy.

2.4 Shocks

There are seven exogenous shocks in the model, five of which (risk premium, TFP, investment-specific technology, government purchases, and monetary policy) follow AR(1) processes:

$$\ln \varepsilon_t^b = \rho_b \ln \varepsilon_{t-1}^b + \eta_t^b \tag{2.12}$$
$$\ln \varepsilon_t^a = \rho_a \ln \varepsilon_{t-1}^a + \eta_t^a \tag{2.13}$$
$$\ln \varepsilon_t^i = \rho_i \ln \varepsilon_{t-1}^i + \eta_t^i \tag{2.14}$$
$$\ln \varepsilon_t^g = \rho_g \ln \varepsilon_{t-1}^g + \rho_{ga}\eta_t^a + \eta_t^g \tag{2.15}$$
$$\ln \varepsilon_t^r = \rho_r \ln \varepsilon_{t-1}^r + \eta_t^r \tag{2.16}$$

and the remaining two - wage and price markup shocks follow ARMA$(1,1)$ processes:

$$\ln \lambda_{w,t} = (1-\rho_w)\ln \lambda_w + \rho_w \ln \lambda_{w,t-1} + \eta_t^w + \mu_w \eta_{t-1}^w \tag{2.17}$$
$$\ln \lambda_{p,t} = (1-\rho_p)\ln \lambda_p + \rho_p \ln \lambda_{p,t-1} + \eta_t^p + \mu_p \eta_{t-1}^p \tag{2.18}$$

3 Parameter Identification

3.1 Model Solution

The economy in the model is assumed to evolve along a deterministic growth path, with γ being the gross rate of growth. To solve the model, we first detrend all growing variables - consumption, investment, capital, real wages, output and government spending, and then all equilibrium conditions are log-linearized around the deterministic steady state of the detrended variables. A detailed discussion of all log-linear equations can be found in Smets and Wouters (2007)

The linearized system can be written in the following way (see the Appendix for details):

$$\Gamma_0 Z_t = \Gamma_1 E_t Z_{t+1} + \Gamma_2 Z_{t-1} + \Gamma_3 U_t \tag{3.1}$$

where Z_t is a 33×1 vector given by $Z_t = \left[Z_t^{\prime f}, Z_t^{\prime s}\right]^{\prime}$, where Z_t^f and Z_t^s are defined as

$$Z_t^f = \left[c_t^f, l_t^f, w_t^f, q_t^f, i_t^f, r_t^{kf}, r_t^f, k_t^f, \bar{k}_{t-1}^f, y_t^f, z_t^f\right]^{\prime}$$

and

$$Z_t^s = \left[c_t^s, l_t^s, \pi_t, w_t^s, q_t^s, i_t^s, r_t^{ks}, r_t^s, k_t^s, \bar{k}_{t-1}^s, y_t^s, z_t^s, \mathrm{mc}_t, \varepsilon_t^b, \varepsilon_t^i, \varepsilon_t^a, \varepsilon_t^g, \varepsilon_t^p, \varepsilon_t^w, \varepsilon_t^r, \eta_t^p, \eta_t^w\right]^{\prime}$$

where the small letter represent the percent deviation of the respective variable from its steady state level[3]. Z^f is a vector collecting the variables in the flexible price and wage version of the economy, and Z^s collects the variables from the sticky price and wage economy. U_t is a vector of the seven structural shocks:

$$U_t = [\eta_t^a, \eta_t^b, \eta_t^I, \eta_t^w, \eta_t^p, \eta_t^g, \eta_t^r]^{\prime}$$

[3] q here is the percent deviation of real value of capital from the steady state level of one.

The coefficient matrices $\Gamma_0,\ \Gamma_1,\ \Gamma_2$ and Γ_3 are functions of a 39×1 vector of deep parameters θ, defined by

$$\begin{aligned}\theta = \quad & [\delta, \lambda_w, g_y, \varepsilon_p, \varepsilon_w, \rho_{ga}, \beta, \mu_w, \mu_p, \alpha, \psi, \varphi, \sigma_c, h, \Phi, \iota_w, \xi_w, \iota_p, \xi_p, \sigma_l, r_\pi, r_{\triangle y}, r_y, \rho, \rho_a, \rho_b, \\ & \rho_g, \rho_I, \rho_r, \rho_p, \rho_w, \gamma, \sigma_a, \sigma_b, \sigma_g, \sigma_I, \sigma_r, \sigma_p, \sigma_w]' \end{aligned} \tag{3.2}$$

There are several algorithms for solving linear rational expectations models like (3.1) (see for instance Blanchard and Kahn (1980), Anderson and Moore (1985), Klein (2000), Christiano (2002), Sims (2002)). Depending on the value of θ, there may exist zero, one, or many stable solutions. Assuming that a unique solution exists, it can be cast in the following form:

$$Z_t = AZ_{t-1} + BU_t \tag{3.3}$$

where A and B are functions of θ, and are unique for each value of θ. We collect the reduced-form parameters in a vector τ, defined as

$$\tau = [\mathbf{vec}(A)',\ \mathbf{vec}(B)']'$$

We also define the function mapping θ into τ as

$$\tau = h(\theta)$$

The deep parameters of the model cannot be estimated directly from (3.3) as some of the variables in Z are not observed. In particular, the observable variables in the model are

$$X_t = \begin{bmatrix} c_t & l_t & \pi_t & w_t & i_t & r_t & y_t \end{bmatrix}$$

and the remaining $39 - 7 = 32$ variables in Z are treated as latent.

Instead, we can write the reduced-form system in a state space form, with transition equation given by (3.3), and the following measurement equation:

$$X_t = CZ_t \tag{3.4}$$

where C is a 7×33 matrix constructed the rows of 33×33 identity matrix.

Assuming that U_t is normally distributed, the conditional log likelihood function $l(X, \theta)$ can be computed recursively using the Kalman filter (see Hamilton (1994, ch.13)).

3.2 Identification of θ

Let Θ be the parameter space of θ. We say that $\theta_0 \in \Theta$ is globally identifiable if no other $\theta_1 \in \Theta$ yields in the same value of the likelihood function. Local identification, on the other hand, requires that the value of the likelihood at θ_0 is unique in some neighborhood of θ_0. Clearly, local identifiability is necessary for a parameter to be globally identified. Finally, when all a priori admissible values $\theta \in \Theta$ are identifiable, we say that the model is identified.

A well-know result from Rothenberg (1971, Theorem 1) is that a necessary and sufficient condition for local identification of θ_0 is that the information matrix, defined by

$$\Im_\theta = -E[l_{\theta\theta}(X,\theta)]$$

has a full rank when evaluated at θ_0. Here $l_{\theta\theta}$ is the Hessian with respect to θ of the log-likelihood function $l(X,\theta)$. Using this condition we can, in principle, determine the identifiability of the model as a whole by evaluating the rank of the information matrix at all points of the parameter space.

The problem with applying this result to determine identifiability in DSGE models is that the mapping from θ to the log-likelihood function is, for most models, not available in analytical form. The likelihood function is determined by A and B, which have to be solved for numerically with some of the algorithms mentioned earlier. This makes it impossible to derive analytically the information matrix by direct differentiation of the log-likelihood function. Using numerical differentiation, on the other hand, is computationally very costly, and is known to be very inaccurate for highly non-linear functions which is typically the case for DSGE models. Not only is the function non-linear, but it has to be evaluated numerically in the first place.

In Iskrev (2007a) we presented an alternative approach for evaluating the information matrix. It is based on a result by Rothenberg (1966) who showed that $\Im_\theta$ can be expressed in the following way[4]

$$\Im_\theta = H'\Im_\tau H \tag{3.5}$$

where $\Im_\tau$ is the information matrix of the unrestricted model, and H is the gradient of h, i.e.

$$H = h_\theta(\theta)$$

Both H and $\Im_\tau$ can be derived analytically. We outline the derivation of H below; see Iskrev (2007a) for references on how $\Im_\tau$ can be computed.

The first step in finding H is to realize that, even though h cannot be written explicitly, we can find an implicit function relating θ and τ. From (3.1) and (3.3) and the law of iterated expectations we obtain the following two sets of equations (see the Appendix for details):

$$(\Gamma_0 - \Gamma_1 A)A - \Gamma_2 = 0 \tag{3.6}$$

$$(\Gamma_0 - \Gamma_1 A)B - \Gamma_3 = 0 \tag{3.7}$$

A and B depend on θ only through τ, while Γ_0, Γ_1, Γ_2 and Γ_3 are functions of θ only. The expressions in (3.6) and (3.7) define an implicit function $F(\theta,\tau(\theta)) = 0$.[5] Therefore, by the implicit function theorem we have[6]

$$H = \frac{\partial \tau(\theta)}{\partial \theta'} = -\big(F_\tau(\theta,\tau(\theta))\big)^{-1} F_\theta(\theta,\tau(\theta)) \tag{3.8}$$

In practice, it is straightforward to compute F_θ and F_τ using standard packages for symbolic calculus, such as the *Symbolic Toolbox* in Matlab. The computation is further simplified by the fact that F can be factored as[7]

$$F(\theta,\tau(\theta)) = F_1(\tau(\theta))F_2(\theta) \tag{3.9}$$

[4]This follows from a straightforward application of the rule for differentiating composite functions.

[5]Evaluating the matrix F proved to be an extremely useful method for detecting and correcting programming errors. See the Appendix for more details on this and a complementary method for doing that.

[6]To apply the implicit function theorem, we need the matrix $F_\tau(\theta,\tau(\theta))$ to be invertible. This was true for all admissible values of θ used in our identification analysis. See below for details.

[7]see the Appendix in Iskrev (2007a)

The approach described above is useful not only because it avoids numerical differentiation, and allows for an accurate evaluation of the information matrix. It may also help in discovering the sources of the identification problems, if such occur. The roots of identification failures may be either in $\mathfrak{I}_\tau$, or H, or both. The first matrix measures how well the reduced form parameters τ are identified, and depends, in part, on the properties of the data, as X is used in its calculation. H, on the other hand, tells us how well identified are the deep parameters θ given τ, and does not depend on the data. Therefore, finding a rank deficient, or poorly conditioned H, means that θ is not identifiable, or is weakly identifiable, due to reasons inherent in the structure of the model.

We should make it clear from the outset that the information matrix approach to identification is for local analysis only. In general, global identification analysis for models that are non-linear in the parameters is not feasible.[8] In Iskrev (2007b) we derive conditions for global identification of the *structural* parameters in linearized DSGE models, i.e. parameters in which the structural equations are linear.[9] However, the goal in the empirical DSGE research is usually to estimate the deep parameters, for which identification can be analyzed only locally.

3.3 Identification analysis procedure

In the previous section we outlined how the information matrix $\mathfrak{I}_\theta$ can be evaluated. Using that approach, we can determine whether the particular value of θ, where the matrix is evaluated, is identifiable or not. The model as a whole is identified if all points from the parameter space Θ are identifiable. It is clearly not feasible to check the rank condition for all points in Θ, and instead we will perform such checks for many randomly drawn points from Θ.[10] Our proposed identification analysis procedure consists of the following steps:

1. Draw randomly a point θ^j from Θ.
2. Check whether the reduced-form solution of the linearized structural model exists and is unique. If both of these conditions are not satisfied, go back to (1).
3. Evaluate the rank and the conditioning of H. If it is of less then full rank, go back to (1).
4. Evaluate the rank and the conditioning of $\mathfrak{I}_\theta$.

In Step (1) we take one a priori admissible value of θ which we then threat as the true parameter value in steps (2) to (4). Upon completion of the procedure, we will know if that value of θ is identifiable, and how strong identification is. Step (2) is necessary to ensure that there exists a unique likelihood function at θ^j. Conditions for existence and uniqueness of the solution can be

[8]See Rothenberg (1971) for more details.

[9]We distinguish between *deep* and *structural* parameters. For instance, if one of the equations in the linearized DSGE model is the New Keynesian Phillips curve

$$\pi_t = \frac{\beta}{1+\varpi\beta}E_t\pi_{t+1} + \frac{(\psi+\nu)(1-\zeta\beta)(1-\zeta)}{(1+\varpi\beta)\zeta}y_t + \frac{\varpi}{1+\varpi\beta}\pi_{t-1} + e_t$$

we call β, ϖ, ψ, ν and ζ deep parameters, and $\gamma_1 = \frac{\beta}{1+\varpi\beta}$, $\gamma_2 = \frac{(\psi+\nu)(1-\zeta\beta)(1-\zeta)}{(1+\varpi\beta)\zeta}$ and $\gamma_3 = \frac{\varpi}{1+\varpi\beta}$ - structural parameters.

[10]Boswijk and Doornik (2003) suggest this approach for checking identification of cointegration relationships.

found in Sims (2002), and are automatically checked by most computer algorithms for solving linear rational expectations models. We call *admissible* the values of θ for which these conditions are satisfied. In Step (3) we check the necessary condition for identification. Finding that H is rank deficient, or poorly conditioned at θ^j, tells us that this particular point of the parameter space is either not identifiable, or is weakly identifiable for structural reasons, i.e. irrespectively of the data. To complete step (4) we need to evaluate $\Im_\tau$, which depends on the data as well as on θ^j. Therefore we need to first generate data Z, assuming that θ^j is the true parameter value. From the rank and conditioning of $\Im_\theta$ we then determine whether θ^j is identified or not, and whether identification, from both the model and the data, is strong or weak.

Before going further we should clarify what we mean by conditioning of a matrix, and how we measure it. The conditioning of a (square) matrix indicates how far the matrix is from being singular. A well conditioned matrix is far from singular, while poorly conditioned one is almost singular. This property of matrices is made precise with the matrix *condition number*, which is the reciprocal of that distance (see section 3.5 for further discussion the condition number). Why is this relevant for us? Remember that the information matrix tells us how informative is the likelihood for the parameters of interest. Thus, the further the information matrix is from singularity, i.e. the smaller is the condition number, the more informative is the likelihood, and better identified are the estimated parameters. When the information matrix is exactly singular, the condition number is equal to infinity, and the likelihood function is absolutely flat in some directions, and is thus completely uninformative with respect to one or more parameters. A poorly conditioned information matrix, on the other hand, has a large condition number, and indicates that the likelihood is nearly flat in some directions, and thus provides very little information for some parameters. This is analogous to collinearity in the linear regression model, and the weak instruments problem - in the instrumental variables setup.

3.4 Identification of the Smets Wouters (2007) model

Now we apply the procedure from the previous section to the model described in section 2. We take the parameter space Θ to be the one defined by the prior distribution of θ, as specified in Smets and Wouters (2007). A summary of that distribution is provided in Table A.1 of Appendix A. This prior distribution is very common in the recent studies using Bayesian methods to estimate similar New Keynesian DSGE models. An alternative approach would be to treat all a priori admissible parameter values as equally likely, that is, to assume uniform priors. The benefit of our approach is that it provides a better coverage of the parts of the space that are considered in the literature as more plausible. For instance, the discount factor β could, theoretically, lie anywhere between 0 and 1. However, values close to .99 are considered to be much more likely than values close to 0. This type of considerations are reflected by the choice of shape and parameters of the prior distribution.

In their estimation procedure Smets and Wouters (2007) treat five deep parameters as known. These are: discount rate δ, share of government spending in GDP g_y, steady state markup in the labor market λ_w, and the two curvature parameters of the aggregation functions in the labor and final good sectors - ϵ_p and ϵ_w[11]. For the first two parameters the reason is that they are difficult

[11] These parameters measure the percent change in the elasticity of demand due to a one percent change in the relative price/wage of the good/labor service, evaluated in steady state. In the simple case, where the aggregator func-

to estimate with the data used in estimation. The markup and the two curvature parameters, on the other hand, are asserted to be unidentifiable. The second claim is easier to check, so we look into it first.

The easiest way to detect lack of identification of one or more deep parameters is to examine matrix $H = \frac{\partial \tau(\theta)}{\partial \theta'}$. It must have full column rank for θ to be identified. Moreover, if a parameter is generally unidentifiable, it would not matter at what admissible value of θ we compute H, as it will be with reduced rank for any $\theta \in \Theta$. In what follows we use the posterior mode of θ reported in Smets and Wouters (2007). When θ includes all 39 parameters listed in (3.2), the rank of H is 36. One of these parameters, however, is the trend growth rate γ for which there is additional information in the trending observed variables that we have not taken into account. Treating γ as known, and computing H for the remaining 38 deep parameters, we conclude that two of them are not identifiable. Closer inspection of H (see section 3.6 for more details) shows us that ϵ_p and ϵ_w are indistinguishable from the Calvo probability parameters ξ_p and ξ_w. In other words, one can identify either ϵ_p or ξ_p but not both simultaneously, and similarly for ϵ_w and ξ_w. No such problem appears to exist regarding λ_w, and when we compute H after ϵ_p and ϵ_w are removed from θ, it has full rank. We conclude, therefore, that there is nothing in the model that makes the wage markup parameter λ_w unidentified. Computing the full information matrix $\mathfrak{I}_\theta = H'\mathfrak{I}_\tau H$ confirms that λ_w is indeed identified at the posterior mode of θ.

As we mentioned above, having γ among the parameters with respect to which H is computed causes additional identification problems. It may be useful to know what the source of these problems is, and whether it would be possible to estimate γ from the stationary version of the model using detrended data. To answer these questions we computed H for θ that includes γ, and sequentially exclude one of the remaining deep parameters. We find that H has reduced rank when δ, β, φ, λ and γ are all included, and is with full rank whenever one of these five parameters is excluded. This implies that γ can be identified, using detrended data only, if either δ, β, φ or λ is kept fixed instead of estimated. This is true, for instance, for the parametrization estimated in Smets and Wouters (2007), where it is assumed that δ is known.

We study the identifiability of the model for six parameterizations that differ in the parameters assumed to be known. The parameters are those assumed known in Smets and Wouters (2007) plus γ. The values of the fixed parameters, reported in table 3.1 below, are also taken from that paper. The trend parameter γ is held fixed in all cases except parametrization 5. In parametrization 1 all other parameters are left free. In parameterizations 2 to 4 one of the other three parameters - δ, λ_w and g_y respectively, is also assumed known. Considering these cases allows us to compare the strength of these parameters' identifiability. In parametrization 5 all parameters except γ are fixed. Since δ is one of them, as we explained above, γ is identified from the stationary model. In parametrization 6 all parameters are assumed known and thus it is closest to the parametrization estimated in Smets and Wouters (2007).[12]

We draw $1,000,000$ points from Θ and perform steps (1) to (3) described in section 3.3 for each one of them. The distributions of the actual draws are shown in Figure A.1. We sort the admissible draws and divide them into 10 groups; then we perform step (4) for 100 points from each group. Thus we compute the full information matrix $\mathfrak{I}_\theta$ for 1,000 admissible points from Θ. We did not

tions $\mathcal{H}$ and $\mathcal{G}$ have the Dixit-Stiglitz functional form, both parameters are equal to zero (see Eichenbaum and Fisher (2007))

[12] The difference is that in Smets and Wouters (2007) γ is estimated using trending data, while in parametrization 6 γ is assumed known.

Table 3.1: Parameterizations

	1	2	3	4	5	6
δ	free	.025	free	free	.025	.025
λ_w	free	free	1.5	free	1.5	1.5
g_y	free	free	free	.18	.18	.18
γ	.431	.431	.431	.431	free	.431

evaluate that matrix for all admissible draws because with the routine we use for evaluation of $\mathfrak{I}_\tau$, it takes very long to compute that matrix.

Between 96% and 98% of the draws were admissible (see table A.2 in Appendix A). There was no stable solution for about .1% to .3% of them, and for about 2% to 4% there were multiple solutions. Matrix H had a full column rank for all of the admissible draws. Thus the necessary condition for identification was satisfied everywhere in the parameter space. Table A.3 in Appendix A reports the ten deciles of the distribution of the condition numbers of H for all six parameterizations. The information matrix $\mathfrak{I}_\theta$ failed to be of full rank for about 2% of the 1000 draws for which it was evaluated. Table A.5 in Appendix A shows the ten deciles of the distribution of the condition numbers of $\mathfrak{I}_\theta$. We see that even though it has a full rank for almost all of the draws, its condition numbers is extremely high which implies that the matrix is poorly conditioned virtually everywhere in the parameter space.

Table 3.2: Cross-correlations

	λ_w	g_y	μ_p	φ	σ_c	h	Φ	ι_w	ξ_w	ι_p	σ_l	$r_{\triangle y}$	r_y
β	.42	**.98**	-.07	.28	.44	-.40	-.85	-.24	-.41	.26	-.36	.42	-.30
φ	**.95**	.26	**-.92**	1	**.90**	-.78	-.24	-.87	**-.95**	**.90**	-.73	**.96**	**-.98**
σ_c	**.99**	.40	-.78	**.90**	1	**-.96**	-.54	**-.91**	**-.99**	.84	-.88	.86	-.88
h	**-.91**	-.36	.69	-.78	**-.96**	1	.56	**.90**	**.92**	-.67	**.95**	-.72	.78
ι_w	**-.90**	-.21	.75	-.87	**-.91**	**.90**	.29	1	**.90**	-.71	.88	-.79	.87
ξ_w	**-.99**	-.38	.84	**-.95**	**-.99**	**.92**	.49	**.90**	1	-.89	.83	**-.91**	**.92**
ι_p	**.90**	.22	-.76	**.90**	.84	-.67	-.33	-.71	-.89	1	-.51	.89	-.81
ξ_p	.52	.89	-.11	.30	.58	-.61	**-.98**	-.34	-.52	.31	-.53	.38	-.31
σ_l	-.82	-.30	.69	-.73	-.88	**.95**	.45	.88	.83	-.51	1	-.66	.75
$r_{\triangle y}$	**.92**	.37	-.84	**.96**	.86	-.72	-.32	-.79	**-.91**	.89	-.66	1	**-.93**
r_y	**-.92**	-.29	.93	**-.98**	-.88	.78	.22	.87	**.92**	-.81	.75	**-.93**	1
ρ	.89	.32	-.74	.88	.84	-.68	-.40	-.69	-.88	**.95**	-.57	**.91**	-.78
ρ_I	.81	.50	-.56	.65	.87	**-.94**	-.68	-.79	-.82	.51	**-.93**	.61	-.68
σ_I	**-.96**	-.45	.81	**-.91**	**-.97**	**.93**	.50	**.91**	**.97**	-.76	**.91**	-.88	**.92**
σ_p	-.81	-.04	**.99**	**-.91**	-.76	.67	0	.74	.82	-.74	.67	-.83	**.92**

Note: Pairwise correlation coefficients $corr(\hat{\theta}_i, \hat{\theta}_j)$ exceeding .95 in absolute value. The values are obtained by inverting and normalizing the information matrix evaluated at θ for which the condition number of the matrix is equal to the median value from Table A.3. High correlation between the estimates of two deep parameters indicates that they are difficult to identify.

The poor conditioning of the information matrix suggests that some of its columns are nearly linearly dependent. Since the information matrix is equal to the inverse of the asymptotic covariance matrix for the estimate of θ, this in turn implies that there exists a strong degree of interdependence among the estimates of some of the deep parameters. This creates identification problems as these parameters' separate effects on the likelihood are difficult to isolate[13]. We can measure the degree of linear dependence by computing the correlations between the columns of the covariance matrix. The complete set of pairwise correlation coefficients may be obtained by inverting and normalizing the information matrix.[14] Table 3.2 shows all pairs of parameters whose estimates have correlation exceeding .95 in absolute value. The correlation coefficients were computed at the value of θ where the condition number of $\mathfrak{I}_\theta$ equals the median of all points at which the information matrix was evaluated. We see, for instance, that the estimate of the wage markup parameter λ_w is extremely highly correlated with ξ_w and σ_c. This partially confirms the claim in Smets and Wouters (2007) that this parameter is difficult to identify in their model, although, as we discussed above, they are mistaken in asserting that λ_w is not identified. Other parameters that would be very difficult to identify at this particular value of θ are σ_c, ξ_w, h and σ_l as well as the policy rule coefficients ρ, ρ_y and $\rho_{\triangle y}$. Although these observations are made on the basis of single point θ from the parameter space, they remained valid for many other parameter values we tried. In addition, as can be seen from Table A.6, very high degree of linear dependence can also be found for other pairs of parameters, such as σ_w, ξ_w, h and λ_2, or r_π, ρ, ρ_I and σ_I. The correlation coefficients reported in Table A.6 were computed at θ equal to the value where the condition number of $\mathfrak{I}_\theta$ equals the 7-th percentile of all points at which the information matrix was evaluated. Since the condition number of matrix is higher - 6.4×10^8 vs. 1.8×10^7, the linear dependencies shown in Table A.6 are substantially stronger than those reported in Table 3.2.

We draw the following three conclusions from this exercise. First, although the necessary and sufficient condition for identification is generally satisfied, the conditioning of the information matrix is very poor, indicating that θ is very weakly identified in most of the parameter space. Second, the reasons for weak identification are mainly in H, which is entirely determined by the structure of the model, and not affected by the data. To see that, remember the relationship between the information matrix $\mathfrak{I}_\theta$ and H (see equation (3.5)). Even when $\mathfrak{I}_\tau$ is very well conditioned, poor conditioning of H will result in poorly conditioned $\mathfrak{I}_\theta$. For instance, suppose that there is very small amount of uncertainty in the estimate of τ, and $\mathfrak{I}_\tau$ has a condition number equal to one. In particular, we let $\mathfrak{I}_\tau = \mathfrak{I}_\tau^*$ be a diagonal matrix whose inverse - the covariance matrix for τ, has non-zero elements equal to 1% of the true values of τ. The deciles of the distribution of the condition numbers of $\mathfrak{I}_\theta = H'\mathfrak{I}_\tau^* H$ are shown in table A.4. If, for instance, the condition number of H is $6e2$ - the median for parametrization 1, we find that the condition number of $\mathfrak{I}_\theta$ is about $3.7e5$. Thus, even though $\mathfrak{I}_\theta$ was computed in relatively small number of points from Θ, our findings regarding H suggest that the identification of θ is generally weak. Third, the strength of identification improves only a little when δ, λ_w and g_y are kept fixed. We see that by comparing the conditioning of H and $\mathfrak{I}_\theta$ for parametrization 1 and 6. The difference is relatively small. Moreover, the improvement seen in parametrization 6 is, at least partly, due to the smaller number of free parameters, and not only because the identifiability of the fixed parameters is much weaker. Of these three parameters, g_y appears to be be the worst identified one. This can be deduced by comparing the conditioning of parametrization 4 with that of parameterizations 2 and 3.

[13] This is easy to see for the linear regression model $y = X\beta + \epsilon$. When two of the regressors, X_i and X_j are nearly collinear, the corresponding coefficients, β_i and β_j will be difficult to identify. Also, since the covariance matrix of the estimate $\hat{\beta}$ is proportionate to $\mathbb{E}X'X$, high collinearity between the regressors implies high correlation between the corresponding elements of $\hat{\beta}$.

[14] That is, we divide each i, j covariance term of the matrix by the product of the standard deviations of variables i and j. Neely, Roy, and Whiteman (2001) also use the correlation matrix of the parameter estimates to determine the sources of identification problems

3.5 Discussion

Our analysis of parameter identification in the Smets and Wouters model suggests that weak identifiability, and not complete failure of identification, is likely to be the more serious problem for DSGE models in general. Even when some parameters are not identifiable, as is the case with ε_p, ξ_p, ε_w, and ξ_w in the model we consider, this is easy to detect - by computing the rank of Jacobian matrix H, and straightforward to deal with - by fixing instead of estimating parameters that lack identification. Unlike identifiability in a strict sense, which is a "either/or" property of the model, what "weak identification" means is harder to define.

In this paper we use the information matrix condition number to measure the strength of identification. As we mentioned above, the condition number of a matrix measures the distance from singularity of the matrix. More precisely, a matrix condition number is the reciprocal of the the distance, in a norm-sense, of a non-singular matrix to the set of singular matrices (Demmel (1987)). Thus, the smaller the condition number of the information matrix, the further it is from singularity, and the stronger the identification of parameters. In the econometrics literature Forchini and Hillier (2005) also propose the condition number of the information matrix as a measure of the strength of identification in parametric models, and show that it is closely related to the concentration parameter, suggested by Stock, Wright, and Yogo (2002) as a measure for the strength of identification in linear models.

The condition number of the information matrix is therefore a natural and simple measure of the strength of parameter identification. It is particularly useful for the purpose for studying identifiability of a model in general, as we do in this paper, as it summarizes in a single number the properties of the whole information matrix, thus allowing us to characterize the strength of identification at a large number of points in the parameter space. It remains unclear, however, how high should the condition number be, to indicate serious identification problems. In the numerical analysis literature it is suggested, as a rule of thumb, that condition numbers greater than the reciprocal of the square root of the machine precisions, lead to unreliable results. For PCs this number is $6.7e7$. On the other hand, matrices with condition numbers greater than $4.5e15$ are numerically singular. With reference to these number, the values reported in Tables A.3 and A.5 suggest very poor identification. The two rule of thumb numbers, however, are solely based on the fact that computers compute with finite precision, and does not take into account the sampling uncertainty that plays important role in estimation with finite data. Thus, the values indicating weak identification are likely to be much lower.[15] What this value is, or how to find such threshold for a particular class of models, is an open question in econometrics in general, and a complete investigation of it is beyond the scope of this paper. Instead, here we first show one possible approach for addressing it, and then discuss the implications for the model we study.

One useful property of the condition number is related to the problem of solving a linear system of equations:

$$Ax = b \tag{3.10}$$

For (3.10) to have a unique solution, A must have a full column rank. It is often important to know how errors in the (estimate of) A and/or b, affect the solution for x. Such information is

[15] I thank James Stock for making this point to me.

provided by the condition number of A, which gives an upper bound on size of the error in x, given errors in A and/or b; that is

$$\frac{\triangle x}{x} \leq cond(A)\left(\frac{\triangle A}{A} + \frac{\triangle b}{b}\right) \tag{3.11}$$

Furthermore, it can be shown that there exist errors $\triangle A$ or $\triangle b$, for which the bound is attained.

Using this property of the condition number, we can find an upper bound for the elements of the asymptotic covariance matrix of the parameter estimates. This follows from the fact that, for the MLE estimator $\hat{\theta}$, the covariance matrix $\mathbf{V}(\hat{\theta})$ is equal to the inverse of the information matrix, and therefore we have

$$\Im_\theta(\hat{\theta})\mathbf{V}(\hat{\theta}) = \mathbf{I} \tag{3.12}$$

This implies (see Edelman and Rao (2005))

$$\frac{\triangle \mathbf{V}(\hat{\theta})}{\mathbf{V}(\hat{\theta})} \leq cond(\Im_\theta(\hat{\theta}))\left(\frac{\triangle \Im_\theta(\hat{\theta})}{\Im_\theta(\hat{\theta})}\right) \tag{3.13}$$

Hence, when $cond(\Im_\theta(\hat{\theta}))$ is large, even small errors in the estimate $\hat{\Im}_\theta(\hat{\theta})$ of $\Im_\theta(\hat{\theta})$, will cause large errors in the estimate of the covariance matrix $\mathbf{V}(\hat{\theta})$. In particular, the standard errors for $\hat{\theta}$ - the diagonal elements of $\mathbf{V}(\hat{\theta})$, could be very imprecisely estimated when $cond(\Im_\theta(\hat{\theta}))$ is high. To see how large these errors could be in our example, we carried out the following Monte Carlo simulation exercise. For each of the ten deciles shown in Table A.5, we assumed that the corresponding matrix $\Im_\theta(\theta)$ is the true information matrix. We then added small errors to the diagonal elements of $\Im_\theta(\theta)$, drawn from standard normal distribution with variance equal to 1% of the true value. The resulting matrix $\tilde{\Im}_\theta(\theta)$ is then inverted and the percentage error in the diagonal elements of $\tilde{\mathbf{V}}(\hat{\theta})$ recorded. Table A.7 in Appendix A shows the results from 1000 repetitions. The reported numbers are percent error in the standard errors of θ for 1 percent error in the corresponding diagonal element of $\Im_\theta$. The results demonstrate that the estimated covariance matrix is very sensitive to even small errors in the estimate of the information matrix, and the higher the condition number of $\Im_\theta$ is, the larger the errors in the estimate of $\mathbf{V}(\hat{\theta})$ tend to be. This shows us that the standard errors obtained by inverting the information matrix are practically meaningless.

An alternative strategy for investigating the relationship between the condition number of the information matrix, and the small sample properties of the estimator, would be to conduct a standard Monte Carlo study, where data is generated and parameters estimated a large number of times. Unfortunately, for a highly non-linear model with a large number of parameters like the one we consider, this approach is impractical. The main problem is that in order to ensure that a global maximum is reached, for even one set of data, is extremely difficult, and requires, at the least, initializing the optimization procedure with a large number of starting points. This makes the time and computational cost of such a study prohibitive. Thus, instead of the main model, we did a Monte Carlo study of small rational expectations model with only two parameters. The model and the simulation results are presented in Appendix C. We can see in Table C.1, a larger condition number is associated with a substantial increase in both the bias and the mean squared error of the maximum likelihood estimator. These results are consistent with the findings in literature on the effects of weak identification in linear models (see for instance Hahn, Hausman, and Kuersteiner (2004) and Flores-Lagunes (2007)).

3.6 Why is identification weak?

The analysis so far suggests that the model as a whole is poorly identified. Moreover, our findings regarding H indicate that the cause for this is in the structure of the linearized model. This is because poor conditioning of H translates into poor conditioning of the information matrix, and consequently, weak identification of θ. In this section we explore this further by focusing on H and trying to understand what makes it ill-conditioned.

To remind the reader, H is the gradient of the function mapping θ into the reduced-form parameters τ. Since the likelihood function depends on θ only through τ (see section 3.2), a necessary condition for identification of θ is that the inverse mapping, from τ to θ, is unique. For this we need H to have full column rank. The condition will fail if, for instance, two of the deep parameters have exactly the same or exactly proportioned effect on τ, i.e. $\frac{\partial\tau}{\partial\theta_1} \propto \frac{\partial\tau}{\partial\theta_2}$. In that case H will have two linearly dependent columns, and will be rank-deficient. This was the case with the curvature parameters ε_w and ε_p, and the Calvo parameters - ξ_w and ξ_p, respectively. Weak-identification, on the other hand, may arise if two deep parameters have very similar, although not exactly the same, or exactly proportioned effect on τ. Then the corresponding two columns of H will be almost but not exactly collinear, and the matrix will be poorly conditioned. In the former case the two parameters are said to be confounded, and in the later - nearly confounded.

To find out whether parameter confoundedness is a serious problem in our model, we computed, for all pairs of parameters θ_i and θ_j, the correlation between $\frac{\partial\tau}{\partial\theta_i}$ and $\frac{\partial\tau}{\partial\theta_j}$. This was done for all admissible values θ that were drawn and used in the analysis described in section 3.3. Thus, unlike the cross-correlations reported in Table 3.2 and Table A.6, which point to troublesome parameters at a single point in the parameter space, here we determine which parameters cause identification problems for the model in general. Figure 1 shows 16 pairs of parameters for which the correlation coefficient exceeded .9 in absolute value for at least 10% of the points. For each pair the figure shows the distribution of the correlation coefficients. For instance, the histogram in the upper-left corner indicates that for about 68% of the points the correlation between $\frac{\partial\tau}{\partial\lambda_w}$ and $\frac{\partial\tau}{\partial\xi_w}$ was .9 or more; for 13% of the points, correlation was between .8 and .9, etc. Similarly, the correlation between $\frac{\partial\tau}{\partial\rho}$ and $\frac{\partial\tau}{\partial r_\pi}$ (upper-right corner) was $-.9$ or less, for about 37% of the points, and between $-.8$ and $-.9$ - for about 33% of the points.

We see that there are three groups of parameters - the preference parameters σ_c, λ, and ρ_b, the labor supply parameters σ_l, λ_w, and ξ_w, and the policy function parameters ρ, r_y, ρ_π, and ρ_r, that, within each group, may have very similar effects on the reduced-form parameters τ. We also find that the same is true, although to a lesser degree, for the capital share parameter α and the discount rate β, as well as the Calvo parameter for prices ξ_p, indexation parameter for prices ι_p and the fixed cost parameter Φ. For the reasons given above these parameters may be hard to identify separately.

We should note, however, that the high degree of near-confoundedness does not occur everywhere in the parameter space. For each of the 16 pairs of parameters in Figure 1, there are points $\theta \in \Theta$ for which the correlations are weak. In addition, there are points for which the sign of the correlation is reversed. For instance, in around 3% of of all draws, the correlation between $\frac{\partial\tau}{\partial\xi_w}$ and $\frac{\partial\tau}{\partial\lambda_w}$ was very close to zero, and in .8% of them it was negative. On the other hand, other pairs of deep parameters may also exhibit near-confoundedness, although not as frequently as those discussed

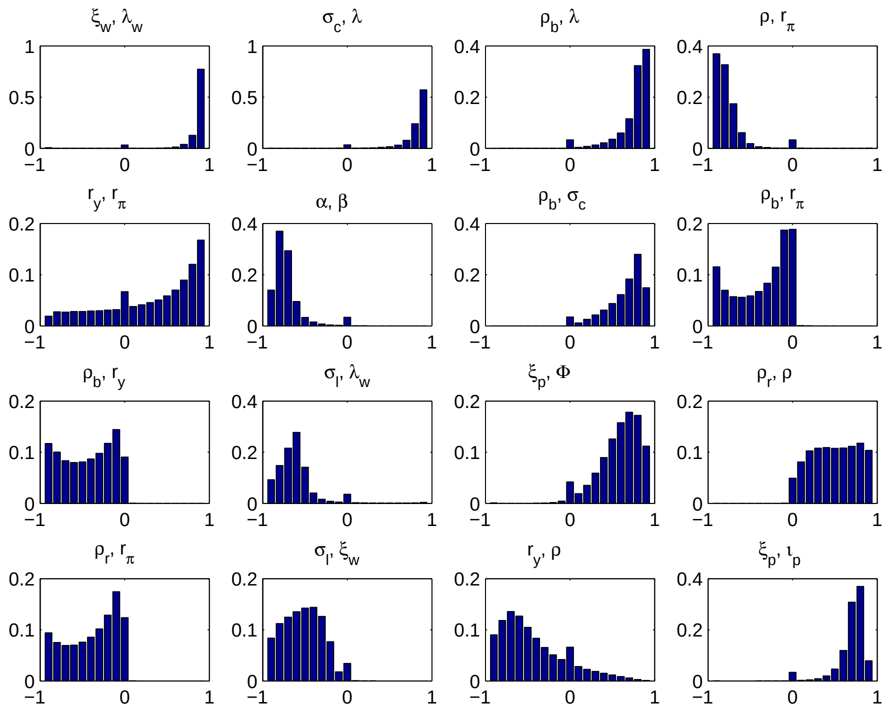

Figure 1:

above.

Remark. It is worth mentioning that parameter near-confoundedness problems in the latest version of the "Smets and Wouters" model are much less severe than those in earlier versions. This is due to the somewhat simpler structure of the current model. For instance, the autocorrelation coefficient of the preference shocks, present in the previous versions of the model, was very difficult to distinguish from the habit persistence and elasticity of intertemporal substitution parameters λ and σ_c. In the current version similar role is played by the risk premium shock, but the near-confoundedness among ρ_b, λ and σ_c is not as strong as before. Another change that improves the identifiability of the model is the simplified monetary policy rule. In the earlier versions of the model the central bank responded to both past inflation and output gap, which was making it difficult to separately identify the response coefficients for current and past inflation and output gap. On the other hand, in the current version the policy shock is assumed to follow an autoregressive process whereas before it was i.i.d. As we can see in Figure 1, ρ_r is potentially difficult to distinguish from ρ and ρ_π.

To summarize, our analysis of parameter identification in the model described in section 2 shows that it is generally weak. The results regarding matrix H indicate that the sources of poor identification are primarily in the structure of the model.[16] One such source is parameter confoundedness which in some cases could be dealt with by fixing the values of some parameters.

[16]Strictly speaking, identification of fully articulated economic models, such as DSGE models, is completely determined by the structure of the model since every aspect of the likelihood can be traced back to the underlying deep parameters and structural relationships. We find it useful, though, to distinguish between the role of H, which depends on θ only, and $\Im_\tau$, which depends on both θ and the data.

Unfortunately this does not seem to provide a fully satisfactory solution of the problem in the model we consider. Parametrization 6, which we found to be relatively better identified, has information matrix that is still relatively poorly conditioned for much of the parameter space. This implies that the data would provide little information for the parameters of this model. Moreover it is unlikely that the identification would be strengthened much by increasing the length or improving the quality of the data used in estimation.

In this section we studied the identification of the theoretical model as it is, without reference to a particular data set used in estimation. Thus the problems we found may arise whenever this or similar DSGE models are estimated. To find out how strong identification is at a particular parameter estimate, that is, conditional on a specific data set, one should examine the conditioning of the information matrix evaluated at that particular point. Furthermore, if Bayesian techniques are used for estimation, in addition to the posterior mode, one could also evaluate the conditioning of the information matrix for all points from the posterior distribution. We return to that in the next section, after the estimation results are presented.

4 Empirical Analysis

The results from the previous section suggest that the likelihood, and therefore the data, is not very informative about the parameters of the model. One consequence of this is that estimating the model using Bayesian techniques, as in Smets and Wouters (2007), one places relatively large weight on the priors compared to the likelihood. To explore this further in this section we estimate the model by maximizing the the likelihood only, and then compare the results with the posterior mode estimates reported in Smets and Wouters (2007)

We start by describing the data to which the model is applied. Then we turn to estimation of the model.

4.1 Data

The model is estimated using quarterly US data over the period 1966:1-2004:4. The observed variables are: real consumption (c), real investment (i), real output (y), real wages (w), hours (h), inflation (π), and the nominal interest rate (r).

Consumption is personal consumption expenditures. Investment is fixed private investment. Wages are hourly compensation for nonfarm business. Real consumption, investment and wages are obtained by deflating the nominal variables with the GDP implicit price deflator. Real output is real GDP. Hours are average hours for nonfarm business. Inflation is the first difference of the log GDP implicit price deflator. Consumption, ivestment, and output are expressed in per capita terms by dividing with civilian population of 16 and older. The nominal interest rate is the quarterly average of the Federal Funds rate.

More details on the definitions and data sources used are provided in the data Appendix to Smets and Wouters (2007).

4.2 Estimation

4.2.1 Maximizing the likelihood function and the posterior density

Both MLE and Bayesian estimation require the evaluation of the likelihood function. To do that we first solve the linearized structural model (3.1) to find the state equation (3.3); then the Kalman filter is used to evaluate the log-likelihood $l(Z;\theta) = \ln L(Z;\theta)$ of the reduced-form model (3.3)-(3.4). In order to keep the estimate of θ within theoretically meaningful bounds, we optimize the likelihood with respect to unbounded variables that are one-to-one transformations of the restricted variables in the θ. The bounds on the parameters in θ are shown in Table D.1, and are the same as those used by Smets and Wouters (2007). In addition, when computing the likelihood we impose the restriction that the model has a unique solution. This is achieved by setting the value of the likelihood to a very small number for values of θ that result in multiple or no solutions.

Using the Bayes rule, the posterior density can be expressed as

$$p(\theta;Z) = \frac{L(Z;\theta)p(\theta)}{p(Z)} \propto L(Z;\theta)p(\theta) \tag{4.1}$$

where $p(\theta)$ denotes the prior distribution of θ. Thus, to maximize the posterior density, we evaluate the likelihood, as before, and the prior $p(\theta)$, which alternatively may be though of as a penalty function.

A well-known practical problem with non-linear optimization/estimation is that one can not be certain that a global maximum is found, and not just a local one. A common strategy for dealing with this is to try many different starting values. Our approach was to combine simulation techniques, gradient and non-gradient based optimization methods. We started with picking ten of the points drawn for the purpose of identification analysis (see section 3.3), which yielded the highest values of the likelihood or the posterior density. Then, taking these points as starting values, we run ten Markov chains generated by the random walk implementation of the Metropolis-Hastings algorithm (we follow Schorfheide (2000), see the appendix for more details). The modes of the distributions generated by each chain were then used as starting values for several optimization routines, and the final maximizer was determined by direct comparison of the resulting values.

4.2.2 Results

We estimate two different parameterizations of the model. In the first one three of the identified parameters - depreciation rate δ, wage markup λ_w, and government spending share in output g_y, are assumed known, and not estimated. This is the parametrization estimated in Smets and Wouters (2007). The values at which these parameters are fixed - .025, 1.5 and .18, respectively, are also taken from that paper. In the second parametrization these parameters are estimated.

We follow Smets and Wouters (2007) and estimate the model using data for the full sample period (1966:1-2004:4), and for two subperiods (1966:1-1979:2 and 1984:1-2004:4). This is done in order to investigate the sources of the differences in the economic environment during these two periods.

The estimation results for the first parametrization are presented in Tables 4.1 (deep parameters), and 4.2 (shock parameters). In addition to our maximum likelihood estimates, and the

Table 4.1: Estimation Results: Deep Parameters

	1966:1-2004:4		1966:1-1979:2		1984:1-2004:4	
Parameter	Bayesian	MLE	Bayesian	MLE	Bayesian	MLE
φ	5.49	7.86	3.62	2.12	6.23	14.98
σ_c	1.40	1.70	1.39	1.21	1.48	1.62
h	0.71	0.70	0.63	0.53	0.69	0.71
ξ_w	0.74	0.88	0.66	0.73	0.75	0.95
σ_l	1.92	2.90	1.52	1.55	2.30	2.32
ξ_p	0.66	0.68	0.56	0.63	0.74	0.79
ι_w	0.59	0.78	0.59	0.86	0.47	0.44
ι_p	0.23	0.01	0.46	0.25	0.21	0.01
ψ	0.55	0.77	0.35	0.16	0.70	1.00
Φ	1.61	1.88	1.43	1.38	1.54	1.63
r_π	2.03	2.49	1.66	3.00	1.77	2.48
ρ	0.82	0.87	0.81	0.91	0.84	0.87
r_y	0.08	0.12	0.18	0.40	0.09	0.07
$r_{\triangle y}$	0.22	0.24	0.21	0.27	0.16	0.19
$\bar{\pi}$	0.82	0.91	0.72	0.76	0.67	0.79
$100(\beta^{-1}-1)$	0.16	0.01	0.15	0.01	0.13	0.03
$\bar{l}$	-0.10	-0.18	0.03	0.04	-0.55	-2.46
γ	0.43	0.42	0.34	0.32	0.45	0.36
α	0.19	0.18	0.20	0.15	0.22	0.19
Log Likelihood:	-840.1	-820.6	-320.2	-303.6	-337.8	-304.4
$cond(\mathfrak{I}_\theta)$:	2.7e7	4.5e7	4.0e7	1.0e9	6.6e7	3.1e8

Note: $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed. $\bar{\pi}$, and $\bar{l}$ are quarterly steady state inflation rate, and steady state hours worked.

posterior mode values from Smets and Wouters (2007), we report the values of the log likelihood as well as the condition number of the information matrix evaluated at the respective point estimates.

Starting with the deep parameters estimated over the whole sample, the results show significant differences between the MLE and Bayesian estimates for most of them. Particularly large is the effect on φ, σ_l, ι_p, ι_w , ξ_w , ψ, and r_π. Smaller, but still substantial are the differences for σ_c, Φ, r_y, $\bar{\pi}$, and $\bar{l}$. For the remaining parameters the estimates are very close.

The maximum likelihood estimates of both Calvo parameters, ξ_p and ξ_w, are higher than their Bayesian estimates. This implies longer average duration of the wage (8.3 vs. 3.9 quarters) and price (3.1 vs. 2.9 quarters) contracts. The estimates of ι_p and ι_w suggest much larger degree of indexation of wages, and much weaker degree of price indexation than those implied by the Bayesian estimates.[17]

The elasticity of the investment adjustment cost function (φ) is also larger according to the ML estimates, as are fixed cost parameter (Φ), and the elasticity of the capacity utilization adjustment cost function (ψ).

[17] This findings are consistent with the remarks in Smets and Wouters (2007) on the effect of relaxing their priors. See their footnote 9.

Overall, for all frictions in the model, except the habit persistence parameter (h), the ML estimates are substantially different and larger than the Bayesian ones. The latter are in turn larger than the respective means of the prior distribution, which is therefore the most likely explanation of the observed discrepancies.

The ML estimate of the monetary policy rule parameters suggest a stronger interest rate response to inflation, output gap, and the change in output gap, as well as higher degree of interest rate smoothing. Again, these differences among the Bayesian and the maximum likelihood estimates can be attributed to the use of the particular prior values.

Table 4.2: Estimation Results: Shock Processes

	1966:1-2004:4		1966:1-1979:2		1984:1-2004:4	
Parameter	Bayesian	MLE	Bayesian	MLE	Bayesian	MLE
ρ_a	0.96	0.97	0.97	0.99	0.94	0.97
ρ_b	0.18	0.14	0.40	0.60	0.14	0.07
ρ_g	0.98	0.98	0.91	0.91	0.97	0.97
ρ_I	0.71	0.70	0.61	0.47	0.65	0.67
ρ_r	0.13	0.01	0.22	0.07	0.30	0.19
ρ_p	0.90	0.95	0.51	0.82	0.75	0.92
ρ_w	0.97	0.98	0.97	1.00	0.83	0.70
ρ_{ga}	0.53	0.45	0.59	0.55	0.40	0.33
μ_w	0.89	0.96	0.85	0.97	0.62	0.61
μ_p	0.74	0.77	0.46	0.98	0.60	0.83
σ_a	0.45	0.43	0.58	0.61	0.35	0.37
σ_b	0.24	0.25	0.23	0.20	0.19	0.20
σ_g	0.52	0.54	0.54	0.52	0.42	0.41
σ_I	0.45	0.45	0.52	0.56	0.40	0.35
σ_r	0.24	0.23	0.20	0.21	0.12	0.12
σ_p	0.14	0.12	0.22	0.26	0.12	0.12
σ_w	0.24	0.27	0.20	0.25	0.22	0.24
Log Likelihood:	-840.1	-821.7	-320.2	-303.6	-337.8	-304.4
$cond(\Im_\theta)$:	2.7e7	4.5e7	4.0e7	1.0e9	6.6e7	3.1e8

Note: $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed

Turning to the estimates of the exogenous shock parameters, presented in Table 4.2, we see that the MLE and Bayesian estimates are quite close. One exception is the autocorrelation parameter of the policy shock (ρ_r), which is estimated to be substantially larger when a prior (with mean of .5) is used. This confirms the observation made in Smets and Wouters (2007) that "the data appear to be very informative on the stochastic processes of for the exogenous disturbances" (p.9). One implication of this is that we should expect that the forecast error variance decompositions of the model variables will be quite similar across the two sets of estimates.

Regarding the estimates obtained using data from the two subsamples, we observe much larger discrepancies among the maximum likelihood and Bayesian estimates of the deep parameters. For some parameters, for instance r_π for the first subperiod, and φ - for the second, the ML estimates were pushed towards the bounds for those parameters. Similar experience, resulting from relaxation of the prior precision, was reported in Onatski and Williams (2004). One possible explanation of these discrepancies is that much less data is used for estimation, which makes the likelihood relative less informative, and the priors - relative more influential with respect to the posterior distribution. This is indicated by the high value of the condition numbers of the information matrix. These values are quite high even when all data is used, but particularly so for two subsample estimates. As we discussed in Section 3.5, having such a poorly conditioned information matrix makes the estimated asymptotic covariance matrix highly sensitive to even small estimation errors in $\mathfrak{I}_\theta$, thus making the estimated standard errors meaningless. Because of that we do not report such errors for the MLE estimates.

Remark. Unlike in Section 3, where we computed condition numbers of the information matrix at the true values of θ, here the parameters are estimated, and therefore subject to sampling uncertainty. Accounting for this uncertainty is straightforward for the estimates obtained with Bayesian methods. We can simply find the posterior distribution of $\mathrm{cond}(\mathfrak{I}_\theta)$. The 5-th and 95-th percentiles of the distribution are $2.4e7$ and $2.9e7$, respectively. It is not obvious how to put similar confidence bounds on the condition number of the information matrix evaluated at the ML estimates.

The results reported in Tables 4.1 and 4.2 were obtained under the assumption that δ, λ_w, and g_y are known and fixed at the values assumed in Smets and Wouters (2007). As we discussed in section 3 the reason given for not estimating these parameters was their poor identification. However, we found evidence supporting that claim only with respect to λ_w. In Table D.2 we report the maximum likelihood estimates of the model parameters obtained when δ, λ_w, and g_y are assumed unknown and also estimated. The values we estimated for these parameters are $\hat{\delta} = .021$, $\hat{\lambda}_w = 1.77$, and $\hat{g}_y = .3$. Turning to the other parameters, the effect is most noticeable for the policy rule parameters, the estimates of all of which increase substantially. The higher condition numbers (6.7e10 vs. 2.7e10) suggest that thee identification of this parametrization is indeed weaker. However, the difference is not particularly large and is, at least partly, due to the large number of parameter estimated in the second case.

Overall, we find that the use of priors have significant effects on the parameter estimates for the model we consider. This by itself does not imply that the model behavior is also affected substantially. To asses the implications of different estimates on the internal dynamics and the propagation mechanism of the model, we next compare the impulse responses to the structural shocks, and the variance decompositions for the observed variables.

4.2.3 Impulse responses and variance decompositions

Impulse responses and variance decompositions are standard tools for gauging the behavior of macroeconomic models, and assessing their credibility. Impulse response analysis allows us to trace the dynamic interactions among economic variables, while the variance decompositions measure the contribution of each structural shock to the total variation of each variable. Here we compare the implications along these two dimensions of three different parameter estimates for the whole sample period (1966:1-2004:4) - the Bayesian and ML estimates for the first parameterizations (columns 2 and 3 of tables 4.1 and 4.2), and the ML estimate for the second parametrization (columns 2 and 6 of Table D.2 in the Appendix). For ease of notation, henceforth we refer to the first two estimates as SW and MLE1, and the the last one - as MLE2.

Figures E.2 - E.8 plot the impulse responses (percent deviations from steady state level) of the seven observed variables (output, consumption, investment, hours, inflation, wages, and interest rate) to a one standard deviation in each of the seven structural shocks (productivity, risk premium, government spending, investment, monetary policy, price and wage markup shocks). Overall, the responses seem reasonable, and are, in most cases, qualitatively similar in the sense of having the same sign on impact and similar dynamics. In particular, most impulse responses implied by the two ML estimates are very close. The most common difference between MLE1 and MLE2 on one hand, and SW - on the other, are in the magnitude and persistence of the responses. For instance, the responses of output and consumption to productivity, investment or price markup shocks, take longer to reach their peaks, and last longer under the MLE, compared to SW estimates. The opposite is true for the response of most variables, and particularly investment and wages, to a wage markup shock. In some cases there is also a substantial difference in the impact effect of the shocks. For instance, wages and inflation respond much more strongly to monetary policy, productivity, risk premium, or government spending shocks, under the SW estimates compared to the MLE ones. In the case of response of wages to exogenous spending shock, the impact effects are also in different directions (see Figure E.4). Under SW the response is positive and remains so for up to 10 quarters, while the two ML estimates imply a smaller and negative response.

Tables F.1 - F.3 report, for the three parameter estimates - SW, MLE1 and MLE2 respectively, the contributions of each structural shock to the forecast error variances of the observed variables at different horizons. As with the impulse responses, the results are broadly similar, with some differences emerging in the medium to long-run horizon. With respect to the determinants of output, for instance, the Bayesian parameter estimates overemphasize, relative to the ML ones, the importance of wage markup, exogenous spending, and risk premium shocks, and underestimate that of sector-neutral productivity, and price markup shocks. Similar differences may be observed regarding inflation. Relative to the ML estimates, the Bayesian estimates overestimate the importance of risk premium, exogenous spending, investment and monetary policy shocks, and underestimate the importance of price markup shocks. These differences are again more significant at medium and long-run horizons.

Similar differences in the importance assigned to different structural shocks can be observed with respect to the other variables in the model. One property that all estimates have in common is that "demand" shocks, such as government spending, risk premium, or investment-specific shocks, are the main driving forces behind the fluctuations in output in a short run. According to both the Bayesian and ML estimates, these shocks 50% to 70% of the forecast error variance of output at horizons of 1 to 4 quarters. On the other hand, at medium to long-run, "supply" shocks - productivity, price and particularly wage markup shocks, are the main driving forces behind the fluctuations in output, explaining between 60% and 80% of the forecast error variance of output at horizons of 10 years and beyond. These observations were made in Smets and Wouters (2007), and as our results show, are robust to the method used for estimation.

5 Conclusion

One of the main promises of the rapidly expanding literature on empirical evaluation of DSGE models, is that we can now estimate micro-founded structural models that until recently had to be calibrated. However, the extent to which this is of practical use depends crucially on whether the estimated parameters are well identified. In this paper we developed a new methodology that can be used to address these questions - are the parameters identified, how strong is identification, are the identification problems inherent in the structure of the model, or due to data deficiencies - for any linearized DSGE model. We then applied this methodology to study the identifiability of parameters in a state-of-the-art monetary DSGE model, that is generally regarded as one of the success stories of the empirical DSGE literature. We found that the parameters of the model are poorly identifiable in much of the parameter space. In addition, our results suggest that the problem is to a large extent inherent in the structure of the model. Thus, it is likely that other models in the empirical DSGE literature, that share features of the model we considered, also suffer from weak parameter identifiability. We showed how parameter confoundedness can be detected and possibly alleviated by reparametrization. In our model this improved, but unfortunately did not fully solved the identification problem. Estimating the model with maximum likelihood methods, we found substantial differences in the parameter estimates compared to those obtained with Bayesian methods. We attribute those differences to the the use of priors in the latter.

Are these differences important? The answer of this question depends on the purpose of estimating the model in the first place. For instance, using estimated DSGE models solely for forecasting purposes does not require knowledge of the values of behavioral or technology parameters. Similarly, if the estimated model is used to conduct impulse response and variance decomposition analysis, then the strength of parameter identification is not that important. We saw evidence to that effect in the last section, where quite different parameter values often implied very similar, and even identical impulse response functions, or variance decomposition results. This should not be surprising, as by definition weak local identification means that different deep parameters imply very similar reduced-form dynamics. However, when estimated DSGE models are used of policy analysis, such as designing optimal monetary policy, the values of the deep parameters may be of crucial importance. This is because for the purpose of such analysis one needs to work with non-linear versions of the model, for which the implications of different parameter values are likely to be stronger than in the linearized version of the model.

Our results may cause one to seriously doubt the validity of parameter estimates reported in some of the empirical DSGE literature. For instance, in their empirical comparison of the US and Euro area business cycles, Smets and Wouters (2005) conclude that the structures of the two economies are very similar, and have not changed much over time. Since the model they estimate is similar to the one in this paper, these findings may be explained with the fact that they use the same prior distributions for both economic areas, and the different sample periods. Of course, if the priors are chosen so that they truly reflect the researcher's a priori beliefs for the parameters of interest, weak identification is not an issue, as long as care is taken to sample from the true posterior distribution. We believe, however, that even when this is the case, conducting and reporting the results of identification analysis as described here, would help in the communication of one's finding to a broader audience, who may not hold the same subjective beliefs as the authors. Providing such information would help the reader assess the relative importance of the data and the priors, and let her judge for herself the credibility of the reported estimates.

Given the increasing popularity of empirical DSGE analysis, one may wonder whether the problems we have discussed in this paper are specific to our model, or endemic, as the analysis in Beyer and Farmer (2004) may lead one to believe. To partially answer this question, we carried out the identification analysis described in section 3.3 for three other DSGE models - a prototypical three-equations New Keynesian model, a standard one-sector stochastic growth model, and a two-country monetary New Open Economy model. The first two are stripped-down versions of our main model, focusing on features that are important in the New Keynesian and the RBC economics, respectively. The third one is an example of a model which is comparable, in terms of size and number of parameters, to our model, but simpler in terms of structural features. More information on the models, and the results from the identification analysis is provided in the Appendix. We find that parameter identification in these models, is much stronger than in the large scale New Keynesian model adopted in this paper. Thus the problem with identification is not generic, and should be addressed for each DSGE model separately.

One way to deal with the identification problems, when such are detected, is to re-parameterize the structural model and estimate parameters that are well identified. This would be an useful approach in situations where the values of the individual deep parameters are not of primary interest, and estimating functions of such parameters is also acceptable. If, for instance, the DSGE model is used for forecasting or to study the dynamic responses of economic variables to structural shocks, this can be accomplished without estimating deep parameters. Moreover, in such situations many of the cross-equation restrictions imposed when the deep parameters are estimated, can be relaxed, thus making the results robust to larger classes of models. Another possible solution is to work with higher order approximations instead of linearized models. McManus (1992) proves that identification failures are much rearer in non-linear than in linear models, and argues that using linear approximations is a major cause for poor parameter identifiability in econometrics. Although the estimation of non-linear DSGE models is computationally much more demanding, recent work by Rubio-Ramirez and Fernandez-Villaverde (2005), An (2005), and Amisano and Tristani (2006) have shown how it could be accomplished. However, the procedures for studying identification proposed here cannot be applied to non-linear models. The development of appropriate methods is left for future work. Another question suggested by the findings in this paper, is whether the difficulties with identification of some of the preference parameters is specific to our model as a whole, or would arise in any model with the same specification of the consumer preferences. This is also left for future investigation.

APPENDICES

A Case Study: Identification

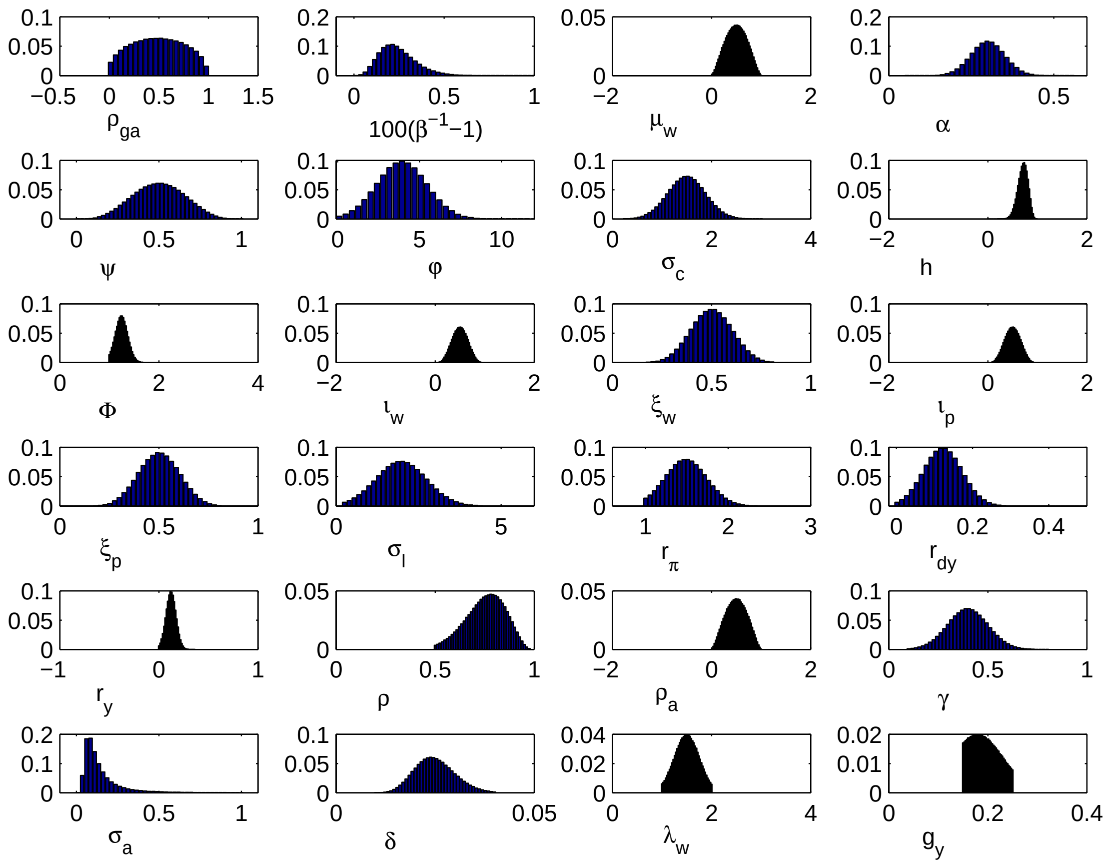

Figure A.1: Distributions of the draws of parameters used in the identification analysis.

Table A.1: Prior Distribution of θ

	Prior		
Parameter	Distr.	Mean	Stdd.
α	$\mathcal{N}$	0.300	0.050
ψ	$\mathcal{B}$	0.500	0.150
φ	$\mathcal{N}$	4.000	1.500
σ_c	$\mathcal{N}$	1.500	0.375
h	$\mathcal{B}$	0.700	0.100
$100(\beta^{-1}-1)$	$\mathcal{G}$	0.250	0.100
Φ	$\mathcal{N}$	1.250	0.125
ι_w	$\mathcal{B}$	0.500	0.150
ξ_w	$\mathcal{B}$	0.500	0.100
ι_p	$\mathcal{B}$	0.500	0.150
ξ_p	$\mathcal{B}$	0.500	0.100
σ_l	$\mathcal{N}$	2.000	0.750
r_π	$\mathcal{N}$	1.500	0.250
$r_{\triangle y}$	$\mathcal{N}$	0.125	0.050
r_y	$\mathcal{N}$	0.125	0.050
ρ	$\mathcal{B}$	0.750	0.100
γ	$\mathcal{N}$	0.400	0.100
δ	$\mathcal{B}$	0.025	0.005
λ_w	$\mathcal{N}$	1.500	0.250
g_y	$\mathcal{N}$	0.180	0.050
ρ_{ga}	$\mathcal{B}$	0.500	0.250
ρ_a	$\mathcal{B}$	0.500	0.200
ρ_b	$\mathcal{B}$	0.500	0.200
ρ_g	$\mathcal{B}$	0.500	0.200
ρ_I	$\mathcal{B}$	0.500	0.200
ρ_r	$\mathcal{B}$	0.500	0.200
ρ_p	$\mathcal{B}$	0.500	0.200
ρ_w	$\mathcal{B}$	0.500	0.200
μ_w	$\mathcal{B}$	0.500	0.200
μ_p	$\mathcal{B}$	0.500	0.200
σ_a	$\mathcal{IG}$	0.100	2.000
σ_b	$\mathcal{IG}$	0.100	2.000
σ_g	$\mathcal{IG}$	0.100	2.000
σ_I	$\mathcal{IG}$	0.100	2.000
σ_r	$\mathcal{IG}$	0.100	2.000
σ_p	$\mathcal{IG}$	0.100	2.000
σ_w	$\mathcal{IG}$	0.100	2.000

Note:$\mathcal{N}$ is Normal distribution, $\mathcal{B}$ is Beta-distribution, $\mathcal{G}$ is Gamma distribution, $\mathcal{IG}$ is Inverse Gamma distribution. The inverse Gamma priors are in the form $p(\sigma;\nu,s) \propto \sigma^{-\nu-1}\exp^{-\nu s^2/2\sigma^2}$; s and ν are given in the Mean column and Stdd. column respectively.

Table A.2: Admissability of draws

Param.	Non-existence	Indeterminacy	Admissible
1	0.30%	3.20%	96.50%
2	0.10%	2.00%	97.90%
3	0.30%	3.10%	96.60%
4	0.20%	3.40%	96.40%
5	0.10%	4.10%	95.80%
6	0.20%	2.40%	97.40%

Note: The total number of draws is $1,000,000$.

Table A.3: Conditioning of H for different parameterizations.

Param.	Decile of cond(H)										
	min	1	2	3	4	5	6	7	8	9	max
1	6.4e1	2.2e2	2.9e2	3.7e2	4.7e2	6.0e2	7.9e2	1.1e3	1.6e3	3.2e3	3.1e11
2	4.8e1	2.0e2	2.8e2	3.6e2	4.5e2	5.8e2	7.6e2	1.0e3	1.6e3	3.1e3	2.9e11
3	4.3e1	1.5e2	1.9e2	2.3e2	2.8e2	3.4e2	4.2e2	5.4e2	7.3e2	1.2e3	2.8e8
4	6.4e1	2.1e2	2.8e2	3.6e2	4.6e2	5.9e2	7.7e2	1.1e3	1.6e3	3.1e3	3.0e11
5	7.0e1	2.8e2	3.9e2	4.9e2	6.1e2	7.4e2	9.1e2	1.1e3	1.5e3	2.1e3	2.8e8
6	3.4e1	1.3e2	1.7e2	2.1e2	2.5e2	3.1e2	3.8e2	4.9e2	6.6e2	1.1e3	2.8e8

Note: $H = \frac{\partial \tau}{\partial \theta}$ is the gradient of the reduced-form parameters w.r.t. θ. $\text{rank}(H) = \dim(\theta)$ is a necessary condition for identification of θ. Large values of cond(H) imply near failure of this condition, thus indicating weak identification. The statistics were computed on the basis of $1,000,000$ random draws of θ.

Table A.4: Conditioning of $H'H$ for different parameterizations.

Param.	Decile of cond($H'H$)										
	min	1	2	3	4	5	6	7	8	9	max
1	4.1e3	4.8e4	8.6e4	1.4e5	2.2e5	3.7e5	6.2e5	1.2e6	2.6e6	1.0e7	9.5e22
2	2.3e3	4.1e4	7.6e4	1.3e5	2.1e5	3.4e5	5.8e5	1.1e6	2.4e6	9.4e6	8.5e22
3	1.8e3	2.2e4	3.6e4	5.5e4	8.0e4	1.2e5	1.8e5	2.9e5	5.3e5	1.4e6	7.6e15
4	4.1e3	4.5e4	8.1e4	1.3e5	2.1e5	3.5e5	6.0e5	1.1e6	2.5e6	9.6e6	9.0e22
5	4.9e3	8.0e4	1.5e5	2.4e5	3.7e5	5.5e5	8.3e5	1.3e6	2.2e6	4.5e6	7.6e15
6	1.2e3	1.6e4	2.7e4	4.2e4	6.3e4	9.4e4	1.4e5	2.4e5	4.4e5	1.1e6	7.6e15

Note: $\text{cond}(\mathfrak{I}_\theta) = \text{cond}(H'H)$ if $\mathfrak{I}_\tau$ is perfectly well conditioned. Thus cond($H'H$) can be thought of as the unattainable lower bound for $\text{cond}(\mathfrak{I}_\theta)$.

Table A.5: Conditioning of $\mathfrak{I}_\theta$ for different parameterizations

Param.	Decile of cond($\mathfrak{I}_\theta$)										
	min	1	2	3	4	5	6	7	8	9	max
1	4.2e5	1.6e6	2.1e6	4.9e6	8.1e6	1.8e7	5.0e7	6.4e8	2.3e9	2.2e10	4.4e24
2	2.7e5	4.7e5	1.3e6	2.9e6	3.3e6	3.7e6	4.9e7	3.5e8	2.2e9	2.1e10	4.1e25
3	1.8e5	1.6e6	1.9e6	2.6e6	4.5e6	1.2e7	1.6e7	4.4e8	1.1e9	2.2e10	1.8e14
4	4.1e5	1.4e6	2.1e6	4.6e6	7.1e6	1.8e7	4.9e7	6.1e8	2.3e9	2.2e10	2.8e24
5	4.3e5	7.6e5	1.3e6	1.5e6	1.8e6	2.0e6	1.5e7	2.8e8	1.0e9	2.1e10	1.6e14
6	1.0e5	4.2e5	1.1e6	1.5e6	1.8e6	2.0e6	1.4e7	1.9e8	1.0e9	2.1e10	1.6e14

Note: $\mathfrak{I}_\theta = H'\mathfrak{I}_\tau H$ is the information matrix for θ. rank($\mathfrak{I}_\theta$) = dim(θ) is a necessary and sufficient condition for identification of θ. Large values of cond($\mathfrak{I}_\theta$) imply near failure of this condition, thus indicating weak identification. These statistics were computed on the basis of $1,000$ random draws of θ.

Table A.6: Cross-correlations

	λ_w	β	μ_p	ψ	σ_c	h	Φ	ξ_w	σ_l	r_π	ρ	ρ_b	ρ_I
α	.77	**.98**	-.54	-.88	.82	-.84	-.75	-.82	**-.92**	**.94**	.87	-.74	.89
ψ	**-.97**	**-.94**	.85	1	**-.98**	**.99**	**.97**	**.98**	**.93**	**-.97**	**-.97**	.59	**-.94**
σ_c	**.99**	.89	-.87	**-.98**	1	-.99	**-.95**	**-.99**	**-.93**	**.96**	**.98**	-.56	**.94**
h	**-.99**	**-.91**	.84	.99	**-.99**	1	**.94**	**.99**	**.95**	**-.97**	**-.97**	.59	**-.93**
ξ_w	**-.99**	-.89	.86	**.98**	**-.99**	**.99**	**.95**	1	**.93**	**-.96**	**-.97**	.58	**-.93**
ξ_p	**.96**	.87	-.85	**-.97**	**.95**	**-.95**	**-.99**	**-.95**	-.83	**.90**	**.92**	-.55	.87
r_π	**.93**	**.97**	-.75	**-.97**	**.96**	**-.97**	-.89	**-.96**	**-.97**	1	**.98**	-.67	**.97**
ρ	**.97**	**.92**	-.84	**-.97**	**.98**	**-.97**	**-.92**	**-.97**	**-.93**	**.98**	1	-.58	**.97**
ρ_I	**.92**	**.92**	-.77	**-.94**	**.94**	**-.93**	-.86	**-.93**	**-.92**	**.97**	**.97**	-.57	1
σ_b	.50	.69	-.24	-.57	.54	-.57	-.48	-.56	-.66	.64	.56	**-.99**	.55
σ_I	**-.91**	**-.94**	.73	.94	**-.94**	**.94**	.84	**.94**	**.97**	**-.99**	**-.97**	.63	**-.99**
σ_p	**-.95**	-.71	**.98**	**.90**	**-.93**	**.90**	**.92**	**.93**	.77	-.82	-.89	.35	-.82
σ_w	**-.99**	-.86	**.90**	**.97**	**-.99**	**.98**	**.94**	**.99**	**.92**	**-.94**	**-.97**	.53	**-.92**

Note: Pairwise correlation coefficients $corr(\hat{\theta}_i, \hat{\theta}_j)$ exceeding .95 in absolute value. The values are obtained by inverting and normalizing the information matrix evaluated at θ for which the condition number of the matrix is equal to the 7-th percentile from Table A.3. High correlation between the estimates of two deep parameters indicates that they are difficult to identify.

Table A.7: Percent error in diag($V(\hat{\theta})$) for 1% error in diag($\Im_\theta(\hat{\theta})$)

Param.	Decile of $\Im_\theta$									
	min	1	2	3	4	5	6	7	8	9
δ	93.6	202.2	-136.8	-294.9	355.7	-96.8	266.6	2245.6	49.3	-236.7
λ_w	102.6	218.2	97.3	211.3	-132.5	-173.0	-265.0	-177.6	62.4	-56.2
g_y	0.4	9.1	-495.3	-56.6	49.6	-587.8	159.5	-23.8	1308.6	1530.8
ρ_{ga}	-19.7	-123.8	-57.8	-34.1	-17.1	22.6	9.2	-16.9	-32.8	-12.3
$\bar{\beta}$	65.5	54.9	-138.4	-736.3	68.6	-225.4	90.2	569.2	-58.5	-89.8
μ_w	-11.3	28.5	2.8	6.7	23.7	-7.5	-25.1	-23.2	-34.7	-324.4
μ_p	18.8	89.3	114.5	-539.5	7.1	-76.6	-68.8	-14.0	128.0	-119.6
α	-39.2	-369.0	69.1	-135.4	-31.2	-116.3	81.2	-27.5	-141.0	-234.6
ψ	52.8	54.0	-64.4	-75.5	-46.8	-13.5	203.6	-30.3	-1471.1	521.7
φ	66.7	65.9	56.0	-98.7	-35.9	-633.6	-164.0	-992.9	291.9	-1407.6
σ_c	-40.2	-47.7	107.7	171.4	-1720.9	-176.7	127.8	203.4	199.3	156.9
h	163.7	-60.5	-42.8	83.9	36.6	1920.0	179.5	61.9	107.0	-136.9
Φ	160.5	-63.2	388.8	509.4	-119.7	1251.6	2346.8	185.1	-113.1	144.4
ι_w	-9.3	6.8	-382.7	-231.4	-34.5	654.0	-123.2	-361.7	109.0	-112.5
ξ_w	319.4	153.1	-1231.1	187.7	59.1	159.9	104.6	327.6	310.5	-150.1
ι_p	99.3	309.5	178.8	549.4	612.1	57.7	-180.6	24.1	-88.5	-69.0
ξ_p	-67.0	59.4	-122.4	172.9	89.2	78.1	-68.2	-78.9	-81.0	-76.3
σ_l	-144.1	134.3	-128.9	1450.5	73.7	30.9	-241.8	41.5	-114.9	123.9
r_π	77.6	-139.4	337.8	102.2	-61.3	-872.6	138.7	-256.9	-9506.0	-4013.2
$r_{\triangle y}$	-118.2	32.0	-48.8	24.9	-86.9	-24.1	-171.3	-4639.8	72.8	-38.5
r_y	71.7	-50.7	143.9	41.8	-216.4	-98.2	84.0	113.2	-198.2	-167.6
ρ	-70.4	108.5	396.0	625.5	-121.6	289.8	-2027.8	97.0	-149.8	-408.1
ρ_a	-36.7	-22.4	772.4	461.4	0.4	-32.5	8.9	-12.5	-29.7	-19.4
ρ_b	-0.8	90.4	-69.7	-233.8	-204.0	-37.4	-90.3	957.3	118.8	179.8
ρ_g	0.8	1.4	0.8	14.5	-4.3	-4.7	38.5	-4.3	38.4	-20.2
ρ_I	-0.6	-87.2	-38.1	2.8	1.4	-179.7	214.4	43.1	12.3	-35.1
ρ_r	6.4	-182.2	5.5	-189.0	1306.1	-78.6	-183.6	-86.1	68.0	11.1
ρ_p	-5.1	1.8	-4.2	748.3	-8.4	-2.9	-227.3	-7.2	-7.7	-21.0
ρ_w	-1.2	-0.4	58.0	-9.5	3.4	2.8	-2.8	63.1	-2.3	136.8
σ_a	-88.9	-70.7	79.8	101.0	39.8	119.0	67.2	17.2	999.0	-348.2
σ_b	-5.9	-70.8	149.8	-148.7	76.2	60.1	268.5	-34.9	173.3	107.3
σ_g	-63.9	14.3	7.5	13.0	34.5	-24.8	-1221.2	14.2	-27.0	-28.7
σ_I	1.0	-15.1	-58.5	-11.4	-27.8	34.8	-234.5	35.2	-14.5	10550.7
σ_r	-167.0	-41.1	-3.4	-191.1	-57.3	1022.8	154.5	-162.7	-646.0	-656.9
σ_p	19.7	50.0	167.8	-227.5	-14.6	117.5	173.0	12.5	42.2	25.4
σ_w	-0.2	-8.6	7.6	-168.3	-12.0	-24.0	-256.9	128.7	25.6	-44.8
cond($\Im_\theta$)	4.2e5	1.6e6	2.1e6	4.9e6	8.1e6	1.8e7	5.0e7	6.4e8	2.3e9	2.2e10

Note:

B Identification: Three Alternative Models

Table B.1: Conditioning of H for 3 different DSGE models

model	Decile										
	min	1	2	3	4	5	6	7	8	9	10
New Keynesian	8.2	22.8	28.1	33.2	38.7	45.1	52.7	62.7	77.1	103.4	8.7e2
RBC	4.9	17.5	23.7	30.1	36.7	44.1	53.5	68.1	95.7	177.7	4.0e8
NOE (2 country)	12.3	41.3	52.2	62.9	75.0	89.6	108.9	137.0	185.4	303.5	1.1e10

Note:$H = \frac{\partial \tau}{\partial \theta}$ is the gradient of the reduced-form parameters w.r.t. θ. $\text{rank}(H) = \dim(\theta)$ is a necessary condition for identification of θ. Large values of $\text{cond}(H)$ imply near failure of this condition, thus indicating weak identification. The statistics were computed on the basis of $1,000,000$ random draws of θ.

C Monte Carlo Study: Small Example

Structural Model

$$\Gamma_0 y_t = \Gamma_1 E_t y_{t+1} + \Gamma_1 y_{t-1} + \Gamma_3 u_t, \tag{C.1}$$

where y is univariate and

$$\Gamma_0 = (1+\delta), \quad \Gamma_1 = (1+\gamma+\gamma^2/2), \quad \Gamma_2 = (\delta-\gamma-\gamma^2/2), \quad \Gamma_3 = e^{\gamma}$$

Parameters: δ, γ.

The **reduced form** solution is:

$$y_t = Ay_{t-1} + Be_t \tag{C.2}$$

where A and B can be calculated by hand:

$$A = \frac{2\delta - 2\gamma + \gamma^2}{2+2\gamma+\gamma^2}, \quad B = \frac{2e^{\gamma}}{2+2\gamma+\gamma^2}$$

Identification problems δ and γ are difficult to identify separately when $\gamma \approx 0$. One way to see that is by computing H given by

$$H = \begin{bmatrix} \frac{\partial A}{\partial \delta} & \frac{\partial A}{\partial \gamma} \\ \frac{\partial B}{\partial \delta} & \frac{\partial B}{\partial \gamma} \end{bmatrix} = \begin{bmatrix} 2/(2+2\gamma+\gamma^2) & -4(1+\gamma)(1+\delta)/(2+2\gamma+\gamma^2)^2 \\ 0 & 2e^{\gamma}\gamma^2/(2+2\gamma+\gamma^2)^2 \end{bmatrix}$$

When $\gamma \approx 0$ the columns of H are almost collinear, which implies that, locally, the effect on A and B of perturbing δ is very similar to that of perturbing γ. Since the likelihood function depends on the parameters only through A and B, this implies that they are poorly identified for $\gamma \approx 0$. For instance, if $\delta = .25$ and $\gamma = .01$, the *condition number* of H is 51247. If $\delta = 3.6$ and $\gamma = 1.4$, on the other hand, the *condition number* of H is 11.

We can also see why the problem arises directly, by realizing that δ and γ only enter the likelihood function as either $f = \frac{1+\gamma+\gamma^2/2}{1+\delta}$ or $g = \frac{e^{\gamma}}{1+\delta}$ (we can write $A = \frac{1-f}{f}$, and $B = \frac{g}{f}$). When $\gamma \approx 0$, f and g are very similar, which make it difficult to separate δ from γ.

Table C.1: Condition number and finite sample properties of MLE: Example

	Relative Bias					Relative MSE				
Parameter	1	2	3	4	5	1	2	3	4	5
δ	-0.3	0.6	1.0	1.0	1.1	1.0	2.7	3.2	3.3	3.5
γ	-0.5	-0.6	1.4	12.5	68.7	0.9	3.9	37.8	376.2	766.8
cond(H)	2.6e1	5.1e2	5.1e4	5.1e6	2.0e7	2.6e1	5.1e2	5.1e4	5.1e6	2.0e7

Note: Results from Monte Carlo study with 1000 repetitions.

D Estimation

Table D.1: Parameter Bounds

Parameter	lower bounds	upper bounds
φ	2.000	15.000
σ_c	0.250	3.000
h	0.001	0.990
ξ_w	0.300	0.950
σ_l	0.250	10.000
ξ_p	0.500	0.950
ι_w	0.010	0.990
ι_p	0.010	0.990
ψ	0.010	1.000
Φ	1.000	3.000
r_π	1.000	3.000
ρ	0.500	0.975
r_y	0.001	0.500
$r_{\triangle y}$	0.001	0.500
$\bar{\pi}$	0.100	2.000
$100(\beta^{-1} - 1)$	0.010	2.000
$\bar{l}$	-10.000	10.000
γ	0.100	0.800
α	0.010	1.000
δ	0.010	0.400
λ_w	1.000	2.000
g_y	0.150	0.300
ρ_a	0.010	1.000
ρ_b	0.010	1.000
ρ_g	0.010	1.000
ρ_I	0.010	1.000
ρ_r	0.010	1.000
ρ_p	0.010	1.000
ρ_w	0.001	1.000
ρ_{ga}	0.010	2.000
μ_w	0.010	1.000
μ_p	0.010	1.000
σ_a	0.010	3.000
σ_b	0.025	5.000
σ_g	0.010	3.000
σ_I	0.010	3.000
σ_r	0.010	3.000
σ_p	0.010	3.000
σ_w	0.010	3.000

Note:

Table D.2: Estimation Results: MLE

Structural				Shock processes			
	1966-2004	1966-1979	1984-2004		1966-2004	1966-1979	1984-2004
φ	7.924	2.000	15.00	ρ_a	0.972	0.993	0.977
σ_c	1.680	1.193	1.617	ρ_b	0.124	0.614	0.092
h	0.724	0.533	0.703	ρ_g	0.976	0.925	0.965
ξ_w	0.853	0.743	0.950	ρ_I	0.690	0.474	0.678
σ_l	2.883	1.606	2.466	ρ_r	0.010	0.069	0.187
ξ_p	0.669	0.637	0.786	ρ_p	0.939	0.827	0.926
ι_w	0.789	0.836	0.438	ρ_w	0.980	0.995	0.656
ι_p	0.010	0.220	0.010	ρ_{ga}	0.510	0.615	0.323
ψ	0.748	0.185	1.000	μ_w	0.967	0.971	0.554
Φ	1.812	1.352	1.627	μ_p	0.755	0.988	0.832
r_π	3.000	3.000	2.600	σ_a	0.433	0.613	0.369
ρ	0.889	0.902	0.878	σ_b	0.257	0.201	0.203
r_y	0.186	0.412	0.090	σ_g	0.514	0.496	0.385
$r_{\triangle y}$	0.275	0.284	0.211	σ_I	0.463	0.575	0.351
$\bar{\pi}$	1.004	1.321	0.782	σ_r	0.228	0.210	0.120
$100(\frac{1}{\beta}-1)$	0.010	0.010	0.016	σ_p	0.122	0.268	0.116
$\bar{l}$	-0.692	-1.414	-2.276	σ_w	0.273	0.245	0.239
γ	0.408	0.318	0.364				
α	0.199	0.168	0.187				
δ	0.021	0.016	0.022				
λ_w	1.768	1.545	1.526				
g_y	0.300	0.300	0.299				
Log L:	-814.1	-301.3	-299.6		-814.1	-301.3	-299.6
$cond(\Im_\theta)$:	6.0e7	1.8e9	4.2e8		6.0e7	1.8e9	4.2e8

Note: δ, λ_w, and g_y are estimated

E Impulse responses

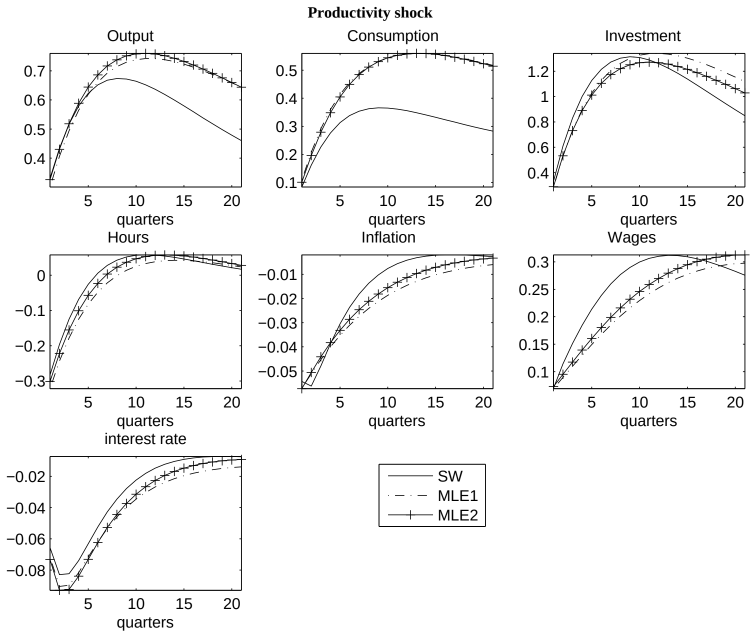

Figure E.2: Impulse Responses to a productivity shock

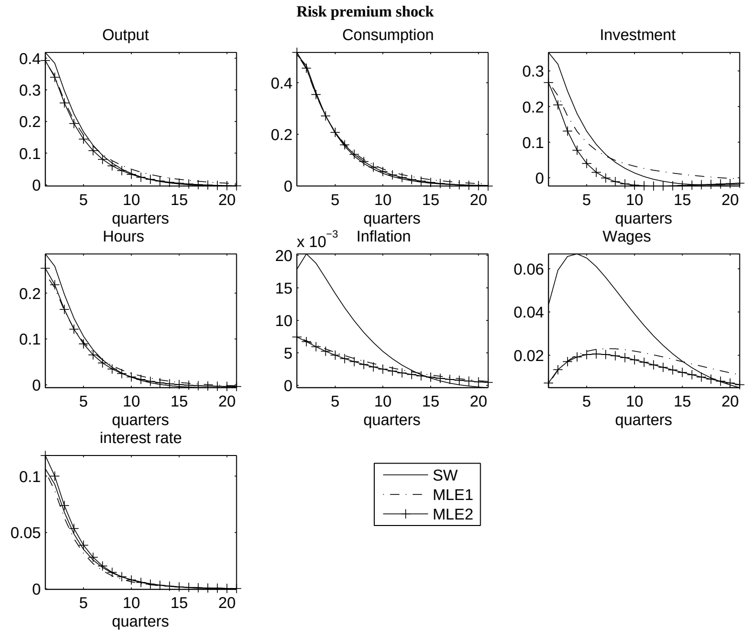

Figure E.3: Impulse Responses to risk premium shock

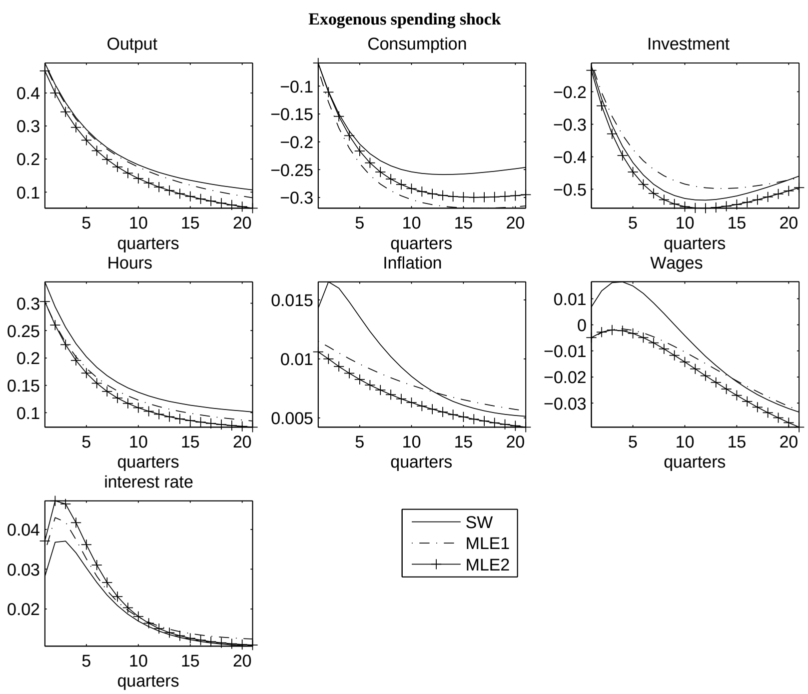

Figure E.4: Impulse Responses to exogenous spending shock

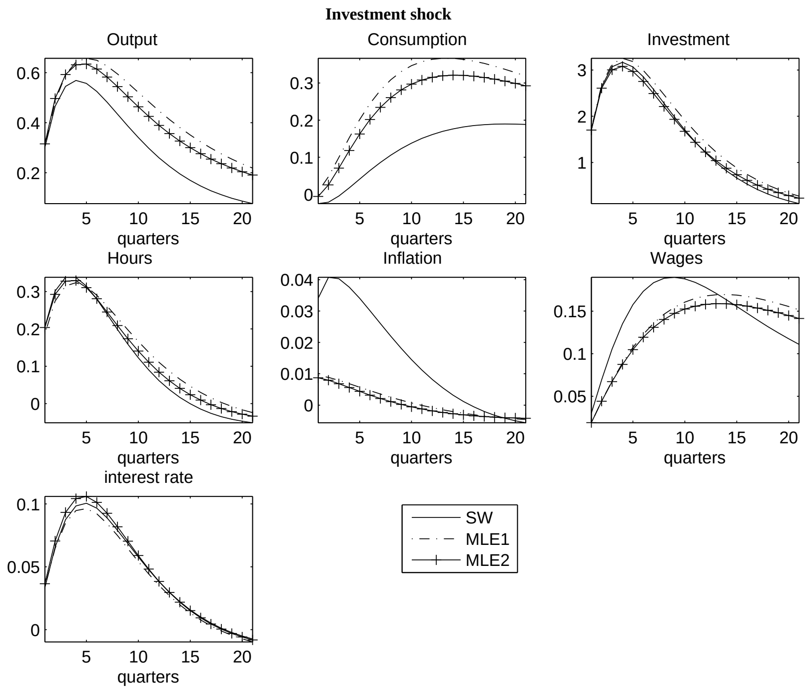

Figure E.5: Impulse Responses to investment shock

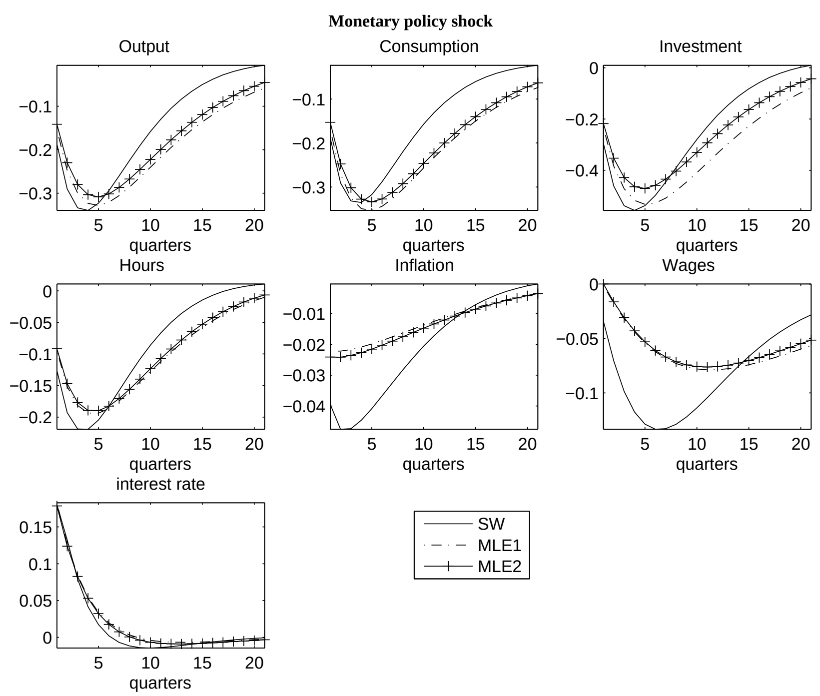

Figure E.6: Impulse Responses to monetary policy shock

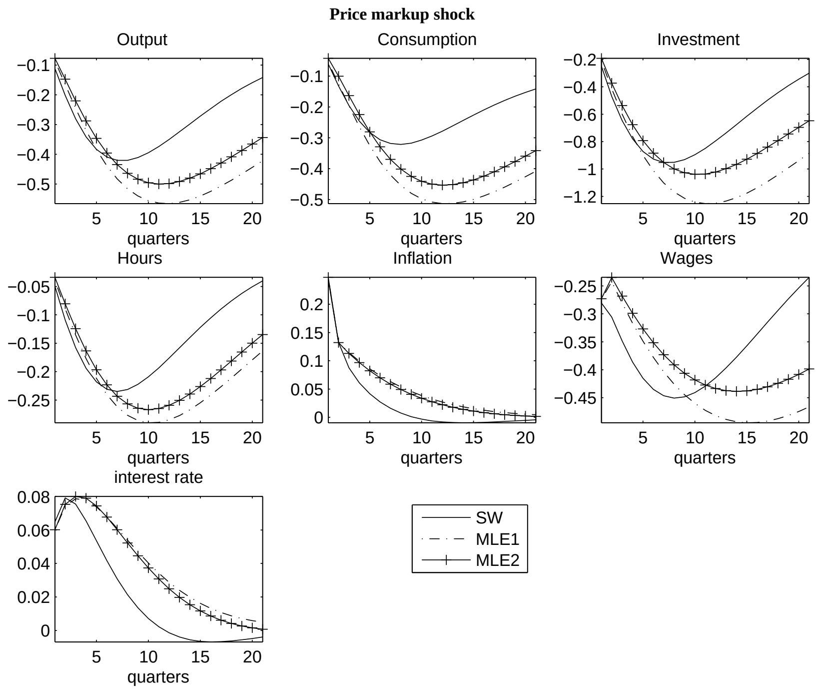

Figure E.7: Impulse Responses to price markup shock

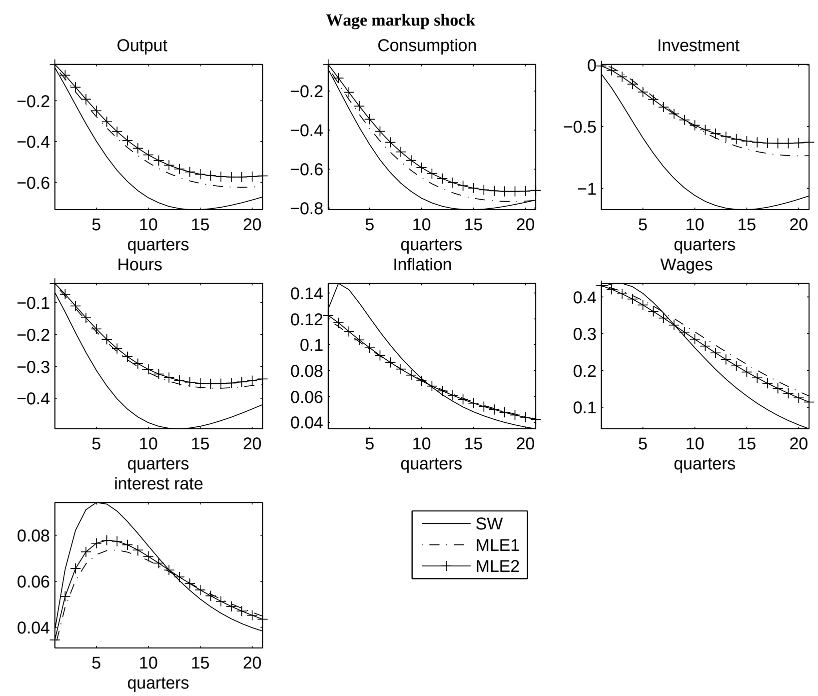

Figure E.8: Impulse Responses to wage markup shock

F Variance Decompositions

Table F.1: Variance Decomposition: Bayesian 1966:1-2004:4

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.164	0.264	0.363	0.136	0.053	0.019	0.002
	Consumption	0.022	0.817	0.011	0.002	0.112	0.012	0.024
	Investment	0.033	0.037	0.004	0.879	0.026	0.019	0.002
	Hours	0.234	0.239	0.337	0.123	0.047	0.007	0.014
	Inflation	0.036	0.004	0.002	0.014	0.019	0.725	0.199
	Wages	0.018	0.007	0.000	0.003	0.004	0.292	0.676
	interest rate	0.077	0.204	0.014	0.020	0.583	0.075	0.027
2	Output	0.193	0.210	0.274	0.200	0.077	0.035	0.011
	Consumption	0.047	0.671	0.022	0.001	0.169	0.030	0.060
	Investment	0.043	0.020	0.006	0.877	0.027	0.025	0.004
	Hours	0.172	0.215	0.292	0.193	0.077	0.021	0.031
	Inflation	0.047	0.006	0.004	0.022	0.030	0.597	0.295
	Wages	0.031	0.009	0.000	0.010	0.011	0.297	0.642
	interest rate	0.106	0.190	0.020	0.052	0.477	0.100	0.055
4	Output	0.242	0.124	0.179	0.250	0.093	0.067	0.044
	Consumption	0.097	0.413	0.043	0.001	0.207	0.072	0.167
	Investment	0.063	0.009	0.008	0.846	0.026	0.038	0.010
	Hours	0.101	0.151	0.231	0.262	0.108	0.056	0.091
	Inflation	0.052	0.007	0.005	0.031	0.042	0.464	0.399
	Wages	0.056	0.011	0.001	0.026	0.022	0.330	0.555
	interest rate	0.132	0.152	0.026	0.129	0.327	0.115	0.118
10	Output	0.312	0.046	0.088	0.196	0.064	0.111	0.183
	Consumption	0.144	0.124	0.064	0.010	0.112	0.111	0.435
	Investment	0.126	0.004	0.019	0.709	0.021	0.069	0.052
	Hours	0.044	0.069	0.145	0.202	0.086	0.114	0.341
	Inflation	0.046	0.007	0.006	0.037	0.053	0.346	0.505
	Wages	0.126	0.008	0.000	0.059	0.032	0.407	0.368
	interest rate	0.119	0.103	0.028	0.224	0.204	0.093	0.230
40	Output	0.308	0.018	0.045	0.089	0.026	0.071	0.443
	Consumption	0.116	0.030	0.078	0.034	0.029	0.052	0.661
	Investment	0.196	0.002	0.045	0.472	0.014	0.069	0.201
	Hours	0.021	0.029	0.096	0.092	0.037	0.067	0.659
	Inflation	0.041	0.006	0.008	0.035	0.048	0.298	0.564
	Wages	0.291	0.004	0.004	0.072	0.022	0.398	0.208
	interest rate	0.105	0.084	0.033	0.206	0.168	0.077	0.327
100	Output	0.295	0.016	0.042	0.079	0.023	0.063	0.482
	Consumption	0.105	0.023	0.090	0.032	0.022	0.042	0.686
	Investment	0.192	0.002	0.051	0.456	0.013	0.066	0.219
	Hours	0.020	0.026	0.105	0.086	0.034	0.062	0.668
	Inflation	0.040	0.006	0.010	0.034	0.046	0.285	0.579
	Wages	0.314	0.004	0.010	0.072	0.021	0.377	0.202
	interest rate	0.104	0.077	0.039	0.194	0.154	0.071	0.361

Note: Based on the posterior mode of θ reported in Smets and Wouters (2007). $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed.

Table F.2: Variance Decomposition: MLE1 1966:1-2004:4

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.150	0.252	0.393	0.151	0.039	0.012	0.002
	Consumption	0.037	0.828	0.015	0.000	0.090	0.010	0.020
	Investment	0.025	0.022	0.004	0.918	0.017	0.014	0.000
	Hours	0.344	0.195	0.304	0.116	0.030	0.005	0.006
	Inflation	0.041	0.001	0.002	0.001	0.006	0.769	0.180
	Wages	0.018	0.000	0.000	0.001	0.000	0.283	0.698
	interest rate	0.091	0.193	0.021	0.021	0.593	0.064	0.017
2	Output	0.183	0.194	0.296	0.235	0.061	0.025	0.007
	Consumption	0.079	0.669	0.031	0.003	0.145	0.027	0.046
	Investment	0.033	0.011	0.005	0.913	0.018	0.019	0.000
	Hours	0.277	0.176	0.273	0.188	0.055	0.016	0.014
	Inflation	0.052	0.001	0.002	0.001	0.009	0.694	0.241
	Wages	0.025	0.000	0.000	0.005	0.001	0.260	0.709
	interest rate	0.130	0.181	0.030	0.052	0.484	0.089	0.033
4	Output	0.234	0.109	0.186	0.312	0.080	0.056	0.024
	Consumption	0.157	0.387	0.056	0.021	0.192	0.074	0.112
	Investment	0.049	0.005	0.007	0.884	0.020	0.033	0.001
	Hours	0.186	0.128	0.224	0.278	0.091	0.052	0.041
	Inflation	0.058	0.001	0.003	0.002	0.012	0.622	0.303
	Wages	0.039	0.001	0.000	0.014	0.003	0.291	0.652
	interest rate	0.162	0.142	0.036	0.125	0.345	0.124	0.067
10	Output	0.312	0.037	0.081	0.276	0.067	0.138	0.090
	Consumption	0.225	0.101	0.073	0.070	0.120	0.161	0.250
	Investment	0.110	0.002	0.015	0.752	0.022	0.090	0.009
	Hours	0.083	0.062	0.145	0.257	0.101	0.177	0.176
	Inflation	0.059	0.001	0.004	0.001	0.015	0.528	0.391
	Wages	0.083	0.001	0.000	0.041	0.010	0.416	0.449
	interest rate	0.156	0.093	0.034	0.207	0.214	0.146	0.149
40	Output	0.370	0.012	0.031	0.130	0.027	0.152	0.278
	Consumption	0.222	0.020	0.085	0.077	0.031	0.133	0.431
	Investment	0.234	0.001	0.040	0.476	0.016	0.156	0.078
	Hours	0.036	0.025	0.089	0.115	0.047	0.178	0.509
	Inflation	0.052	0.001	0.006	0.003	0.015	0.432	0.491
	Wages	0.219	0.001	0.004	0.057	0.009	0.540	0.171
	interest rate	0.142	0.071	0.039	0.186	0.166	0.124	0.271
100	Output	0.371	0.010	0.026	0.110	0.022	0.130	0.330
	Consumption	0.215	0.015	0.104	0.064	0.023	0.105	0.473
	Investment	0.240	0.001	0.049	0.453	0.015	0.149	0.094
	Hours	0.039	0.022	0.097	0.105	0.041	0.158	0.539
	Inflation	0.052	0.001	0.008	0.004	0.014	0.406	0.516
	Wages	0.275	0.001	0.011	0.055	0.008	0.504	0.147
	interest rate	0.145	0.063	0.048	0.172	0.146	0.116	0.311

Note: $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed.

Table F.3: Variance Decomposition: MLE2 1966:1-2004:4

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.176	0.255	0.361	0.165	0.033	0.010	0.001
	Consumption	0.032	0.863	0.011	0.000	0.075	0.006	0.013
	Investment	0.027	0.023	0.006	0.918	0.015	0.012	0.000
	Hours	0.303	0.215	0.306	0.138	0.028	0.004	0.005
	Inflation	0.041	0.001	0.001	0.001	0.007	0.762	0.187
	Wages	0.020	0.000	0.000	0.001	0.000	0.280	0.698
	interest rate	0.091	0.239	0.023	0.023	0.542	0.061	0.020
2	Output	0.209	0.194	0.271	0.249	0.052	0.020	0.004
	Consumption	0.073	0.723	0.024	0.001	0.128	0.018	0.033
	Investment	0.035	0.011	0.007	0.914	0.016	0.017	0.000
	Hours	0.240	0.192	0.273	0.217	0.051	0.013	0.012
	Inflation	0.051	0.001	0.002	0.001	0.010	0.685	0.250
	Wages	0.028	0.000	0.000	0.005	0.001	0.255	0.712
	interest rate	0.129	0.222	0.033	0.058	0.435	0.086	0.037
4	Output	0.265	0.109	0.170	0.320	0.071	0.046	0.018
	Consumption	0.162	0.441	0.049	0.013	0.185	0.058	0.092
	Investment	0.053	0.004	0.011	0.884	0.018	0.029	0.001
	Hours	0.158	0.139	0.224	0.310	0.088	0.045	0.037
	Inflation	0.056	0.001	0.002	0.001	0.014	0.613	0.313
	Wages	0.046	0.001	0.000	0.014	0.003	0.279	0.657
	interest rate	0.157	0.172	0.040	0.138	0.303	0.117	0.073
10	Output	0.359	0.038	0.073	0.268	0.062	0.119	0.081
	Consumption	0.254	0.117	0.071	0.057	0.123	0.143	0.235
	Investment	0.120	0.002	0.023	0.749	0.019	0.077	0.011
	Hours	0.074	0.069	0.146	0.271	0.103	0.163	0.175
	Inflation	0.055	0.001	0.003	0.001	0.018	0.518	0.404
	Wages	0.103	0.001	0.000	0.041	0.010	0.389	0.456
	interest rate	0.147	0.114	0.038	0.228	0.187	0.134	0.153
40	Output	0.426	0.012	0.027	0.126	0.025	0.123	0.260
	Consumption	0.252	0.023	0.084	0.073	0.031	0.110	0.426
	Investment	0.252	0.001	0.057	0.488	0.014	0.119	0.070
	Hours	0.036	0.028	0.087	0.123	0.049	0.158	0.520
	Inflation	0.047	0.001	0.004	0.002	0.017	0.423	0.505
	Wages	0.279	0.000	0.006	0.060	0.009	0.468	0.178
	interest rate	0.130	0.091	0.040	0.207	0.151	0.114	0.267
100	Output	0.424	0.010	0.022	0.106	0.021	0.103	0.314
	Consumption	0.248	0.017	0.097	0.062	0.023	0.083	0.472
	Investment	0.262	0.001	0.066	0.459	0.013	0.113	0.086
	Hours	0.039	0.024	0.089	0.109	0.042	0.138	0.559
	Inflation	0.045	0.001	0.005	0.003	0.016	0.394	0.537
	Wages	0.350	0.000	0.015	0.058	0.008	0.418	0.151
	interest rate	0.131	0.081	0.045	0.191	0.135	0.105	0.310

Note: δ, λ_w, and g_y are estimated.

Table F.4: Variance Decomposition: Bayesian 1966:1-1979:2

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.279	0.265	0.300	0.091	0.049	0.013	0.002
	Consumption	0.086	0.792	0.000	0.001	0.110	0.011	0.000
	Investment	0.061	0.083	0.020	0.794	0.034	0.008	0.000
	Hours	0.141	0.315	0.370	0.111	0.057	0.005	0.001
	Inflation	0.072	0.011	0.004	0.008	0.022	0.644	0.240
	Wages	0.106	0.023	0.002	0.003	0.009	0.359	0.499
	interest rate	0.114	0.342	0.038	0.018	0.345	0.068	0.075
2	Output	0.336	0.240	0.218	0.119	0.068	0.019	0.001
	Consumption	0.155	0.676	0.001	0.001	0.143	0.021	0.003
	Investment	0.088	0.057	0.028	0.779	0.036	0.010	0.002
	Hours	0.091	0.328	0.312	0.161	0.091	0.013	0.004
	Inflation	0.097	0.017	0.007	0.014	0.038	0.427	0.401
	Wages	0.204	0.030	0.003	0.006	0.016	0.283	0.458
	interest rate	0.129	0.338	0.046	0.040	0.247	0.063	0.139
4	Output	0.457	0.161	0.140	0.129	0.076	0.024	0.012
	Consumption	0.312	0.457	0.005	0.001	0.158	0.034	0.033
	Investment	0.154	0.031	0.044	0.713	0.035	0.011	0.012
	Hours	0.055	0.279	0.259	0.212	0.128	0.027	0.039
	Inflation	0.092	0.020	0.008	0.019	0.053	0.257	0.551
	Wages	0.382	0.029	0.003	0.012	0.025	0.185	0.363
	interest rate	0.132	0.280	0.051	0.079	0.144	0.041	0.274
10	Output	0.622	0.068	0.064	0.078	0.045	0.015	0.109
	Consumption	0.503	0.173	0.013	0.000	0.081	0.020	0.209
	Investment	0.329	0.013	0.079	0.468	0.022	0.007	0.081
	Hours	0.054	0.165	0.170	0.163	0.101	0.024	0.322
	Inflation	0.066	0.015	0.008	0.018	0.055	0.181	0.657
	Wages	0.682	0.014	0.001	0.015	0.020	0.080	0.187
	interest rate	0.096	0.180	0.044	0.098	0.086	0.026	0.471
40	Output	0.659	0.032	0.031	0.037	0.021	0.007	0.212
	Consumption	0.553	0.059	0.014	0.008	0.028	0.007	0.330
	Investment	0.528	0.006	0.066	0.223	0.010	0.003	0.163
	Hours	0.043	0.092	0.100	0.096	0.057	0.014	0.599
	Inflation	0.056	0.012	0.007	0.015	0.044	0.146	0.720
	Wages	0.874	0.005	0.002	0.010	0.008	0.030	0.071
	interest rate	0.080	0.132	0.034	0.076	0.064	0.019	0.594
100	Output	0.670	0.029	0.028	0.034	0.019	0.006	0.212
	Consumption	0.600	0.044	0.013	0.009	0.021	0.005	0.308
	Investment	0.563	0.006	0.059	0.199	0.009	0.003	0.160
	Hours	0.078	0.086	0.095	0.091	0.053	0.013	0.584
	Inflation	0.060	0.012	0.007	0.015	0.042	0.140	0.725
	Wages	0.895	0.004	0.002	0.009	0.006	0.024	0.060
	interest rate	0.098	0.122	0.033	0.071	0.059	0.018	0.600

Note: Based on the posterior mode of θ reported in Smets and Wouters (2007). $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed.

Table F.5: Variance Decomposition: MLE1 1966:1-1979:2

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.403	0.252	0.225	0.037	0.068	0.014	0.001
	Consumption	0.267	0.562	0.011	0.005	0.120	0.032	0.004
	Investment	0.039	0.166	0.041	0.660	0.084	0.001	0.008
	Hours	0.102	0.380	0.345	0.057	0.102	0.013	0.001
	Inflation	0.160	0.006	0.001	0.000	0.033	0.421	0.379
	Wages	0.188	0.007	0.000	0.000	0.001	0.121	0.683
	interest rate	0.089	0.527	0.030	0.004	0.299	0.000	0.050
2	Output	0.473	0.239	0.152	0.040	0.085	0.009	0.001
	Consumption	0.371	0.449	0.015	0.005	0.123	0.026	0.012
	Investment	0.065	0.150	0.058	0.616	0.102	0.001	0.007
	Hours	0.057	0.424	0.277	0.072	0.151	0.015	0.004
	Inflation	0.151	0.006	0.001	0.001	0.039	0.392	0.411
	Wages	0.306	0.013	0.000	0.000	0.005	0.108	0.567
	interest rate	0.089	0.589	0.032	0.009	0.194	0.008	0.078
4	Output	0.603	0.167	0.093	0.034	0.091	0.005	0.006
	Consumption	0.543	0.274	0.018	0.005	0.107	0.013	0.039
	Investment	0.134	0.116	0.096	0.516	0.123	0.010	0.004
	Hours	0.041	0.396	0.228	0.081	0.213	0.011	0.029
	Inflation	0.118	0.006	0.001	0.001	0.043	0.419	0.413
	Wages	0.465	0.016	0.001	0.001	0.013	0.129	0.376
	interest rate	0.084	0.570	0.032	0.016	0.119	0.038	0.140
10	Output	0.745	0.068	0.040	0.016	0.054	0.027	0.050
	Consumption	0.698	0.099	0.015	0.003	0.050	0.015	0.121
	Investment	0.336	0.052	0.153	0.274	0.100	0.072	0.014
	Hours	0.058	0.245	0.153	0.057	0.186	0.076	0.225
	Inflation	0.091	0.005	0.001	0.000	0.052	0.342	0.509
	Wages	0.724	0.008	0.000	0.001	0.017	0.101	0.150
	interest rate	0.068	0.453	0.028	0.020	0.085	0.053	0.293
40	Output	0.787	0.025	0.015	0.006	0.020	0.015	0.130
	Consumption	0.746	0.027	0.007	0.001	0.014	0.006	0.198
	Investment	0.636	0.019	0.100	0.099	0.037	0.046	0.064
	Hours	0.063	0.109	0.072	0.026	0.083	0.050	0.596
	Inflation	0.064	0.003	0.000	0.000	0.039	0.242	0.652
	Wages	0.928	0.002	0.000	0.000	0.005	0.032	0.033
	interest rate	0.045	0.297	0.019	0.013	0.057	0.037	0.531
100	Output	0.710	0.012	0.007	0.003	0.010	0.007	0.250
	Consumption	0.669	0.009	0.003	0.000	0.004	0.002	0.313
	Investment	0.749	0.008	0.045	0.044	0.017	0.021	0.116
	Hours	0.123	0.038	0.026	0.009	0.029	0.018	0.757
	Inflation	0.030	0.002	0.000	0.000	0.018	0.112	0.839
	Wages	0.971	0.001	0.000	0.000	0.002	0.013	0.013
	interest rate	0.024	0.125	0.008	0.006	0.024	0.016	0.797

Note: $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed.

Table F.6: Variance Decomposition: MLE2 1966:1-1979:2

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.434	0.250	0.190	0.047	0.062	0.015	0.001
	Consumption	0.275	0.558	0.014	0.007	0.109	0.034	0.004
	Investment	0.032	0.180	0.038	0.666	0.078	0.002	0.005
	Hours	0.069	0.413	0.319	0.079	0.103	0.016	0.001
	Inflation	0.176	0.006	0.001	0.001	0.028	0.410	0.378
	Wages	0.191	0.006	0.000	0.000	0.001	0.118	0.685
	interest rate	0.095	0.547	0.028	0.007	0.270	0.000	0.052
2	Output	0.498	0.238	0.127	0.050	0.078	0.009	0.001
	Consumption	0.380	0.445	0.018	0.008	0.111	0.027	0.012
	Investment	0.054	0.166	0.055	0.625	0.094	0.001	0.004
	Hours	0.035	0.452	0.248	0.095	0.148	0.016	0.005
	Inflation	0.163	0.006	0.001	0.001	0.032	0.400	0.398
	Wages	0.307	0.011	0.000	0.001	0.004	0.106	0.571
	interest rate	0.092	0.606	0.029	0.013	0.172	0.009	0.078
4	Output	0.621	0.167	0.076	0.041	0.082	0.006	0.006
	Consumption	0.552	0.272	0.022	0.007	0.095	0.014	0.038
	Investment	0.115	0.133	0.093	0.532	0.117	0.008	0.002
	Hours	0.035	0.418	0.199	0.102	0.203	0.012	0.031
	Inflation	0.129	0.006	0.001	0.001	0.035	0.429	0.400
	Wages	0.467	0.013	0.000	0.001	0.010	0.127	0.381
	interest rate	0.084	0.588	0.027	0.022	0.104	0.041	0.134
10	Output	0.758	0.068	0.033	0.019	0.049	0.027	0.046
	Consumption	0.710	0.099	0.019	0.003	0.044	0.015	0.110
	Investment	0.300	0.064	0.156	0.291	0.097	0.071	0.021
	Hours	0.074	0.256	0.133	0.069	0.174	0.076	0.219
	Inflation	0.102	0.005	0.001	0.000	0.043	0.358	0.492
	Wages	0.725	0.007	0.000	0.001	0.013	0.102	0.151
	interest rate	0.069	0.480	0.024	0.025	0.076	0.056	0.271
40	Output	0.807	0.026	0.013	0.007	0.019	0.015	0.113
	Consumption	0.773	0.027	0.010	0.002	0.012	0.006	0.170
	Investment	0.582	0.024	0.112	0.110	0.038	0.048	0.085
	Hours	0.068	0.123	0.070	0.034	0.084	0.054	0.567
	Inflation	0.073	0.003	0.000	0.000	0.033	0.259	0.631
	Wages	0.929	0.001	0.000	0.000	0.004	0.032	0.032
	interest rate	0.049	0.335	0.017	0.018	0.055	0.041	0.486
100	Output	0.794	0.013	0.007	0.004	0.010	0.008	0.165
	Consumption	0.772	0.009	0.004	0.001	0.004	0.002	0.208
	Investment	0.738	0.011	0.051	0.049	0.017	0.021	0.113
	Hours	0.165	0.057	0.033	0.016	0.039	0.025	0.665
	Inflation	0.044	0.002	0.000	0.000	0.019	0.154	0.780
	Wages	0.972	0.001	0.000	0.000	0.001	0.013	0.013
	interest rate	0.034	0.189	0.010	0.010	0.031	0.023	0.702

Note: δ, λ_w, and g_y are estimated.

Table F.7: Variance Decomposition: Bayesian 1984:1-2004:4

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.082	0.237	0.390	0.222	0.057	0.011	0.000
	Consumption	0.005	0.844	0.006	0.001	0.129	0.010	0.005
	Investment	0.019	0.024	0.004	0.918	0.026	0.008	0.001
	Hours	0.271	0.187	0.314	0.174	0.044	0.005	0.005
	Inflation	0.027	0.002	0.005	0.014	0.024	0.756	0.172
	Wages	0.002	0.004	0.001	0.009	0.006	0.127	0.851
	interest rate	0.119	0.192	0.036	0.052	0.476	0.088	0.035
2	Output	0.095	0.182	0.294	0.317	0.090	0.021	0.001
	Consumption	0.013	0.703	0.014	0.008	0.218	0.025	0.018
	Investment	0.025	0.013	0.006	0.913	0.029	0.011	0.002
	Hours	0.210	0.162	0.269	0.258	0.077	0.012	0.012
	Inflation	0.040	0.003	0.008	0.022	0.039	0.617	0.270
	Wages	0.003	0.005	0.001	0.017	0.012	0.108	0.854
	interest rate	0.142	0.153	0.045	0.103	0.403	0.092	0.062
4	Output	0.124	0.109	0.196	0.396	0.124	0.038	0.013
	Consumption	0.038	0.451	0.032	0.033	0.316	0.060	0.070
	Investment	0.041	0.007	0.010	0.885	0.034	0.016	0.007
	Hours	0.139	0.114	0.215	0.341	0.122	0.029	0.039
	Inflation	0.049	0.005	0.012	0.032	0.060	0.470	0.371
	Wages	0.006	0.006	0.002	0.035	0.023	0.100	0.830
	interest rate	0.160	0.110	0.054	0.201	0.279	0.083	0.113
10	Output	0.208	0.050	0.110	0.364	0.121	0.057	0.089
	Consumption	0.092	0.166	0.072	0.091	0.245	0.083	0.251
	Investment	0.098	0.003	0.027	0.769	0.039	0.026	0.038
	Hours	0.077	0.067	0.162	0.319	0.142	0.056	0.178
	Inflation	0.052	0.005	0.018	0.038	0.085	0.380	0.422
	Wages	0.025	0.006	0.002	0.079	0.045	0.112	0.733
	interest rate	0.153	0.075	0.061	0.299	0.176	0.057	0.179
40	Output	0.324	0.032	0.075	0.270	0.085	0.044	0.170
	Consumption	0.150	0.070	0.170	0.150	0.117	0.045	0.299
	Investment	0.167	0.003	0.077	0.620	0.033	0.023	0.077
	Hours	0.062	0.052	0.160	0.265	0.118	0.050	0.293
	Inflation	0.053	0.005	0.026	0.047	0.087	0.371	0.410
	Wages	0.109	0.005	0.006	0.117	0.052	0.113	0.598
	interest rate	0.152	0.067	0.085	0.306	0.161	0.053	0.175
100	Output	0.331	0.031	0.075	0.268	0.084	0.044	0.168
	Consumption	0.153	0.063	0.209	0.154	0.105	0.041	0.274
	Investment	0.167	0.003	0.082	0.617	0.033	0.023	0.076
	Hours	0.063	0.051	0.170	0.264	0.116	0.049	0.288
	Inflation	0.054	0.005	0.030	0.049	0.087	0.367	0.407
	Wages	0.117	0.005	0.012	0.118	0.051	0.111	0.587
	interest rate	0.154	0.065	0.099	0.305	0.155	0.052	0.171

Note: Based on the posterior mode of θ reported in Smets and Wouters (2007). $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed.

Table F.8: Variance Decomposition: MLE1 1984:1-2004:4

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.103	0.222	0.394	0.234	0.032	0.011	0.005
	Consumption	0.027	0.832	0.004	0.017	0.088	0.016	0.017
	Investment	0.009	0.006	0.002	0.968	0.007	0.009	0.000
	Hours	0.340	0.163	0.290	0.172	0.023	0.008	0.003
	Inflation	0.031	0.000	0.001	0.001	0.001	0.819	0.147
	Wages	0.004	0.000	0.000	0.001	0.000	0.110	0.885
	interest rate	0.085	0.190	0.027	0.031	0.537	0.116	0.015
2	Output	0.115	0.156	0.290	0.356	0.050	0.021	0.011
	Consumption	0.058	0.651	0.008	0.049	0.154	0.038	0.042
	Investment	0.012	0.003	0.002	0.963	0.008	0.012	0.000
	Hours	0.274	0.137	0.256	0.261	0.044	0.018	0.009
	Inflation	0.045	0.000	0.002	0.002	0.002	0.729	0.221
	Wages	0.006	0.000	0.000	0.001	0.000	0.093	0.900
	interest rate	0.118	0.150	0.037	0.066	0.481	0.120	0.028
4	Output	0.135	0.083	0.177	0.468	0.069	0.042	0.025
	Consumption	0.113	0.352	0.015	0.131	0.213	0.080	0.096
	Investment	0.019	0.001	0.003	0.945	0.010	0.020	0.001
	Hours	0.182	0.094	0.204	0.369	0.077	0.046	0.029
	Inflation	0.058	0.000	0.003	0.003	0.002	0.633	0.300
	Wages	0.010	0.000	0.000	0.003	0.000	0.099	0.888
	interest rate	0.151	0.106	0.046	0.138	0.369	0.129	0.059
10	Output	0.181	0.029	0.078	0.470	0.072	0.090	0.080
	Consumption	0.163	0.086	0.020	0.250	0.153	0.130	0.197
	Investment	0.051	0.001	0.007	0.863	0.016	0.052	0.010
	Hours	0.078	0.046	0.130	0.388	0.107	0.127	0.124
	Inflation	0.071	0.000	0.005	0.007	0.003	0.526	0.388
	Wages	0.029	0.000	0.000	0.011	0.000	0.155	0.804
	interest rate	0.166	0.067	0.051	0.215	0.239	0.127	0.134
40	Output	0.296	0.010	0.031	0.302	0.045	0.101	0.214
	Consumption	0.230	0.019	0.028	0.245	0.061	0.102	0.315
	Investment	0.170	0.001	0.022	0.624	0.018	0.096	0.069
	Hours	0.042	0.022	0.088	0.233	0.080	0.154	0.381
	Inflation	0.080	0.000	0.009	0.013	0.003	0.474	0.421
	Wages	0.157	0.000	0.002	0.058	0.002	0.281	0.501
	interest rate	0.176	0.053	0.064	0.220	0.190	0.106	0.191
100	Output	0.325	0.009	0.029	0.281	0.041	0.096	0.219
	Consumption	0.260	0.017	0.035	0.225	0.055	0.092	0.315
	Investment	0.180	0.001	0.024	0.611	0.018	0.096	0.071
	Hours	0.059	0.021	0.090	0.234	0.074	0.157	0.365
	Inflation	0.081	0.000	0.010	0.013	0.003	0.473	0.420
	Wages	0.238	0.000	0.005	0.064	0.003	0.259	0.430
	interest rate	0.191	0.050	0.070	0.226	0.178	0.099	0.187

Note: $\delta = .025$, $\lambda_w = 1.5$ and $g_y = .18$ are fixed.

Table F.9: Variance Decomposition: MLE2 1984:1-2004:4

qrt		product-ivity	risk premium	exog. spend.	invest-ment	monetary policy	price markup	wage markup
1	Output	0.118	0.237	0.371	0.228	0.032	0.010	0.004
	Consumption	0.041	0.834	0.003	0.010	0.084	0.013	0.014
	Investment	0.008	0.005	0.002	0.970	0.006	0.008	0.000
	Hours	0.345	0.176	0.276	0.169	0.024	0.007	0.003
	Inflation	0.031	0.000	0.001	0.001	0.001	0.815	0.150
	Wages	0.004	0.000	0.000	0.001	0.000	0.111	0.884
	interest rate	0.090	0.225	0.030	0.032	0.497	0.111	0.015
2	Output	0.135	0.168	0.271	0.348	0.050	0.019	0.009
	Consumption	0.086	0.661	0.007	0.033	0.148	0.031	0.035
	Investment	0.011	0.003	0.002	0.965	0.007	0.011	0.000
	Hours	0.277	0.151	0.244	0.259	0.045	0.017	0.008
	Inflation	0.045	0.000	0.002	0.002	0.001	0.725	0.225
	Wages	0.006	0.000	0.000	0.001	0.000	0.091	0.901
	interest rate	0.121	0.180	0.041	0.069	0.444	0.116	0.029
4	Output	0.621	0.167	0.076	0.041	0.082	0.006	0.006
	Consumption	0.552	0.272	0.022	0.007	0.095	0.014	0.038
	Investment	0.115	0.133	0.093	0.532	0.117	0.008	0.002
	Hours	0.035	0.418	0.199	0.102	0.203	0.012	0.031
	Inflation	0.129	0.006	0.001	0.001	0.035	0.429	0.400
	Wages	0.467	0.013	0.000	0.001	0.010	0.127	0.381
	interest rate	0.084	0.588	0.027	0.022	0.104	0.041	0.134
10	Output	0.215	0.030	0.070	0.460	0.070	0.083	0.072
	Consumption	0.227	0.091	0.017	0.216	0.150	0.118	0.181
	Investment	0.049	0.001	0.008	0.873	0.014	0.048	0.007
	Hours	0.080	0.051	0.125	0.392	0.110	0.124	0.119
	Inflation	0.071	0.000	0.003	0.008	0.002	0.523	0.392
	Wages	0.032	0.000	0.000	0.010	0.000	0.154	0.804
	interest rate	0.157	0.080	0.053	0.230	0.220	0.124	0.136
40	Output	0.352	0.010	0.026	0.291	0.042	0.091	0.188
	Consumption	0.307	0.020	0.022	0.226	0.058	0.089	0.278
	Investment	0.179	0.001	0.022	0.635	0.017	0.089	0.057
	Hours	0.045	0.025	0.082	0.239	0.084	0.153	0.372
	Inflation	0.080	0.000	0.006	0.014	0.003	0.470	0.427
	Wages	0.177	0.000	0.002	0.053	0.002	0.275	0.491
	interest rate	0.166	0.064	0.061	0.231	0.176	0.104	0.197
100	Output	0.395	0.009	0.024	0.264	0.038	0.084	0.186
	Consumption	0.358	0.017	0.026	0.203	0.050	0.078	0.268
	Investment	0.197	0.001	0.023	0.617	0.016	0.088	0.058
	Hours	0.069	0.023	0.081	0.239	0.077	0.156	0.354
	Inflation	0.081	0.000	0.006	0.014	0.003	0.469	0.427
	Wages	0.284	0.000	0.004	0.058	0.002	0.246	0.406
	interest rate	0.186	0.060	0.064	0.234	0.165	0.098	0.192

Note: δ, λ_w, and g_y are estimated.

References

Altig, D., L. Christiano, M. Eichenbaum, and J. Linde (2005): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," NBER Working Papers 11034, National Bureau of Economic Research, Inc, available at http://ideas.repec.org/p/nbr/nberwo/11034.html.

Amisano, G., and O. Tristani (2006): "Euro area inflation persistence in an estimated nonlinear," Computing in economics and finance 2006, Society for Computational Economics, available at http://ideas.repec.org/p/sce/scecfa/347.html.

An, A. (2005): "Bayesian Estimation of DSGE Models: Lessons from Second Order Approximations," .

An, S., and F. Schorfheide (2005): "Bayesian Analysis of DSGE Models," CEPR Discussion Papers 5207, C.E.P.R. Discussion Papers, available at http://ideas.repec.org/p/cpr/ceprdp/5207.html.

Anderson, G., and G. Moore (1985): "A linear algebraic procedure for solving linear perfect foresight models," *Economics Letters*, 17(3), 247–252, available at http://ideas.repec.org/a/eee/ecolet/v17y1985i3p247-252.html.

Beyer, A., and R. E. A. Farmer (2004): "On the indeterminacy of New-Keynesian economics," Working Paper Series 323, European Central Bank, available at http://ideas.repec.org/p/ecb/ecbwps/20040323.html.

Blanchard, O. J., and C. M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48(5), 1305–11, available at http://ideas.repec.org/a/ecm/emetrp/v48y1980i5p1305-11.html.

Boivin, J., and M. Giannoni (2006): "DSGE Models in a Data-Rich Environment," NBER Working Papers 12772, National Bureau of Economic Research, Inc, available at http://ideas.repec.org/p/nbr/nberwo/12772.html.

Boswijk, H. P., and J. Doornik (2003): "Identifying, Estimating and Testing Restricted Cointegrated Systems: An Overview," Economics Papers 2003-W10, Economics Group, Nuffield College, University of Oxford, available at http://ideas.repec.org/p/nuf/econwp/0310.html.

Canova, F., and L. Sala (2006): "Back to square one: identification issues in DSGE models," Working Paper Series 583, European Central Bank, available at http://ideas.repec.org/p/ecb/ecbwps/20060583.html.

Christiano, L., M. Eichenbaum, and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45, available at http://ideas.repec.org/a/ucp/jpolec/v113y2005i1p1-45.html.

Christiano, L. J. (2002): "Solving dynamic equilibrium models by a method of undetermined coefficients," *Computational Economics*, 20(1-2).

DelNegro, M., F. Schorfheide, F. Smets, and R. Wouters (2005): "On the fit and forecasting performance of New-Keynesian models," Working Paper Series 491, European Central Bank, available at http://ideas.repec.org/p/ecb/ecbwps/20050491.html.

Demmel, J. (1987): "The condition number and the distance to the nearest ill-posed problem," *Numerische Mathematik*, 51(251289).

Edelman, A., and R. Rao (2005): "Random Matrix Theory," *Acta Numerica*, 1(65).

Eichenbaum, M., and J. Fisher (2007): "Estimating the Frequency of Reoptimisation in Calvo-style Models," *Journal of Monetary Economics, forthcoming.*

Flores-Lagunes, A. (2007): "Finite sample evidence of IV estimators under weak instruments," *Journal of Applied Econometrics*, 22(3), 677–694, available at http://ideas.repec.org/a/jae/japmet/v22y2007i3p677-694.html.

Forchini, G., and G. Hillier (2005): "Ill-conditioned problems, Fisher information and weak instruments," CeMMAP working papers CWP04/05, Centre for Microdata Methods and Practice, Institute for Fiscal Studies, available at http://ideas.repec.org/p/ifs/cemmap/04-05.html.

Greenwood, J., Z. Hercowitz, and P. Krusell (2000): "The role of investment-specific technological change in the business cycle," *European Economic Review*, 44(1), 91–115, available at http://ideas.repec.org/a/eee/eecrev/v44y2000i1p91-115.html.

Hahn, J., J. Hausman, and G. Kuersteiner (2004): "Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations," *Econometrics Journal*, 7(1), 272–306, available at http://ideas.repec.org/a/ect/emjrnl/v7y2004i1p272-306.html.

Hamilton, J. (1994): *Time Series Analysis*. Princeton University Press.

Hansen, L. P., E. R. McGrattan, and T. J. Sargent (1994): "Mechanics of forming and estimating dynamic linear economies," Staff Report 182, Federal Reserve Bank of Minneapolis, available at http://ideas.repec.org/p/fip/fedmsr/182.html.

Iskrev, N. (2007a): "Evaluating the information matrix in linearized DSGE models," .

——— (2007b): "Identification and Estimation of DSGE model - An Integrated Approach," .

Juillard, M., P. Karam, D. Laxton, and P. Pesenti (2006): "Welfare-based monetary policy rules in an estimated DSGE model of the US economy," Working Paper Series 613, European Central Bank, available at http://ideas.repec.org/p/ecb/ecbwps/20060613.html.

Justiniano, A., and G. E. Primiceri (2006): "The Time Varying Volatility of Macroeconomic Fluctuations," NBER Working Papers 12022, National Bureau of Economic Research, Inc, available at http://ideas.repec.org/p/nbr/nberwo/12022.html.

Klein, P. (2000): "Using the generalized Schur form to solve a multivariate linear rational expectations model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423, available at http://ideas.repec.org/a/eee/dyncon/v24y2000i10p1405-1423.html.

Levin, A. T., A. Onatski, J. C. Williams, and N. Williams (2005): "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models," NBER Working Papers 11523, National Bureau of Economic Research, Inc, available at http://ideas.repec.org/p/nbr/nberwo/11523.html.

McManus, D. A. (1992): "How common is identification in parametric models?," *Journal of Econometrics*, 53(1-3), 5–23, available at http://ideas.repec.org/a/eee/econom/v53y1992i1-3p5-23.html.

Neely, C. J., A. Roy, and C. H. Whiteman (2001): "Risk Aversion versus Intertemporal Substitution: A Case Study of Identification Failure in the Intertemporal Consumption Capital Asset Pricing Model," *Journal of Business & Economic Statistics*, 19(4), 395–403, available at http://ideas.repec.org/a/bes/jnlbes/v19y2001i4p395-403.html.

Onatski, A., and N. Williams (2004): "Empirical and policy performance of a forward-looking monetary model," *Proceedings*, (Mar), available at http://ideas.repec.org/a/fip/fedfpr/y2004imarx6.html.

Pesaran, M. H. (1989): *The Limits to Rational Expectations*. Basil Blackwell, Oxford.

Rotemberg, J. J., and M. Woodford (1998): "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy: Expanded Version," NBER Technical Working Papers 0233, National Bureau of Economic Research, Inc, available at http://ideas.repec.org/p/nbr/nberte/0233.html.

Rothenberg, T. J. (1966): "Efficient Estimation with a priori Information: A Classical Approach," Discussion paper.

Rothenberg, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39(3), 577–91, available at http://ideas.repec.org/a/ecm/emetrp/v39y1971i3p577-91.html.

Rubio-Ramirez, J. F., and J. Fernandez-Villaverde (2005): "Estimating dynamic equilibrium economies: linear versus nonlinear likelihood," *Journal of Applied Econometrics*, 20(7), 891–910, available at http://ideas.repec.org/a/jae/japmet/v20y2005i7p891-910.html.

Sargent, T. J. (1976): "The Observational Equivalence of Natural and Unnatural Rate Theories of Macroeconomics," *Journal of Political Economy*, 84(3), 631–40, available at http://ideas.repec.org/a/ucp/jpolec/v84y1976i3p631-40.html.

Schmitt-Grohe, S., and M. Uribe (2004): "Optimal Operational Monetary Policy in the Christiano-Eichenbaum-Evans Model of the U.S. Business Cycle," NBER Working Papers 10724, National Bureau of Economic Research, Inc, available at http://ideas.repec.org/p/nbr/nberwo/10724.html.

Schorfheide, F. (2000): "Loss function-based evaluation of DSGE models," *Journal of Applied Econometrics*, 15(6), 645–670, available at http://ideas.repec.org/a/jae/japmet/v15y2000i6p645-670.html.

——— (2007): *The New Palgrave Dictionary of Economics*chap. Bayesian Methods in Macroeconometrics. Palgrave Macmillan.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20, available at http://ideas.repec.org/a/kap/compec/v20y2002i1-2p1-20.html.

Smets, F., and R. Wouters (2005): "Comparing shocks and frictions in US and euro area business cycles: a Bayesian DSGE Approach," *Journal of Applied Econometrics*, 20(2), 161–183, available at http://ideas.repec.org/a/jae/japmet/v20y2005i2p161-183.html.

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *The American Economic Review*, 97(3), 586–606.

Stock, J. H., J. H. Wright, and M. Yogo (2002): "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20(4), 518–29, available at http://ideas.repec.org/a/bes/jnlbes/v20y2002i4p518-29.html.