University of São Paulo

School of Economics, Business Administration and Accounting at

Ribeirão Preto

Department of Economics

Postgraduate Program in Economics

Estimation and Identification of a DSGE model: an Application of
the Data Cloning Methodology

Pedro Luiz Paulino Chaim

Supervisor: Márcio Poletti Laurini

Ribeirão Preto

2016

Prof. Dr. Marco Antonio Zago
Dean of the University of São Paulo

Prof. Dr. Dante Pinheiro Martinelli
Dean of the School of Economics, Business Administration and Accounting
at Ribeirão Preto

Prof. Dr. Renato Leite Marcondes
Head of the Department of Economics

Prof. Dr. Cláudio Ribeiro de Lucinda
Director of the Postgraduate Program in Economics

PEDRO LUIZ PAULINO CHAIM

Estimation and Identification of a DSGE model: an Application
of the Data Cloning Methodology

Master thesis presented to the Postgraduate Program in Economics of the School of Economics, Business Administration and Accounting at Ribeirão Preto of the University of São Paulo, as a requirement for the title of Master of Sciences.

Supervisor: Márcio Poletti Laurini

Ribeirão Preto

2016

I authorize the reproduction and total or partial disclosure of this work, by any conventional or electronic mean, for purpose of research, as long as the source is indicated.

Chaim, Pedro Luiz Paulino

Estimation and Identification of a DSGE model: an Application of the Data Cloning Methodology. Ribeirão Preto, 2016.

55 p. : il. ; 30 cm

Master Thesis, presented to the School of Economics, Business Administration and Accounting at Ribeirão Preto/USP Postgraduate Program in Economics.

Supervisor: Márcio Poletti Laurini.

1. DSGE. 2. Estimation. 3. Bayesian Asymptotic Theory. 4. Maximum Likelihood.

Acknowledgements

Agradeço primeiramente à minha família que tornou possível que eu me focasse nos meus estudos. Também aos novos amigos que fiz, que tornaram o curso tão mais agradável.

Ao professor Márcio Poletti Laurini, pela excelente orientação e valiosas dicas. Também agradeço aos professores Fábio Augusto Reis Gomes e Jefferson Donizeti Pereira Bertolai, pelos comentários na minha qualificação e pré-defesa, mas principalmente pelas valiosas lições desde o início do curso.

A todos os professores pelas aulas que tive ao longo do mestrado e que são os responsáveis pelo conhecimento que agreguei ao longo desses dois anos de mestrado.

Abstract

CHAIM, P. L. P. **Estimation and Identification of a DSGE model: an Application of the Data Cloning Methodology**. 2016. Master Thesis - School of Economics, Business Administration and Accounting at Ribeirão Preto, University of São Paulo, Ribeirão Preto 2016.

We apply the data cloning method developed by Lele et al. (2007) to estimate the model of Smets and Wouters (2007). The data cloning algorithm is a numerical method that employs replicas of the original sample to approximate the maximum likelihood estimator as the limit of Bayesian simulation-based estimators. We also analyze the identification properties of the model. We measure the individual identification strength of each parameter by observing the posterior volatility of data cloning estimates, and access the identification problem globally through the maximum eigenvalue of the posterior data cloning covariance matrix. Our results indicate that the model is only poorly identified. The system displays bad global identification properties, and most of its parameters seem locally ill-identified.

Key Words: DSGE. Estimation. Identification. Data Cloning. Bayesian Asymptotics. Maximum Likelihood.

Resumo

CHAIM, P. L. P. **Estimação e Identificação de um Modelo DSGE: uma Aplicação da Metodologia *Data Cloning***. 2016. Dissertação (Mestrado) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto 2016.

Neste trabalho aplicamos o método *data cloning* de Lele et al. (2007) para estimar o modelo de Smets e Wouters (2007). O algoritmo *data cloning* é um método numérico que utiliza réplicas da amostra original para aproximar o estimador de máxima verossimilhança como limite de estimadores Bayesianos obtidos por simulação. Nós também analisamos a identificação dos parâmetros do modelo. Medimos a identificação de cada parâmetro individualmente ao observar a volatilidade *a posteriori* dos estimadores de *data cloning*. O maior autovalor da matriz de covariância *a posteriori* proporciona uma medida global de identificação do modelo. Nossos resultados indicam que o modelo de Smets e Wouters (2007) não é bem identificado. O modelo não apresenta boas propriedades globais de identificação, e muitos de seus parâmetros são localmente mal identificados.

Palavras-chaves: DSGE. Estimação. Identificação. Data Cloning. Estatística Bayesiana. Máxima Verossimilhança.

List of Figures

Figure 1 – Single sample MCMC estimate posterior distributions. (1)	43
Figure 2 – Single sample MCMC estimate posterior distributions. (2)	43
Figure 3 – Single sample MCMC estimate posterior distributions. (3)	44
Figure 4 – Single sample MCMC estimate posterior distributions. (4)	44
Figure 5 – 25 Clones MCMC estimate posterior distributions. (1)	45
Figure 6 – 25 Clones MCMC estimate posterior distributions. (2)	45
Figure 7 – 25 Clones MCMC estimate posterior distributions. (3)	46
Figure 8 – 25 Clones MCMC estimate posterior distributions. (4)	46
Figure 9 – Normalized standard error. (1)	48
Figure 10 – Normalized standard error. (2)	48
Figure 11 – Normalized standard error. (3)	49
Figure 12 – Normalized standard error. (4)	49
Figure 13 – Standardized maximum eigenvalue of the posterior covariance matrix .	50

List of Tables

Table 1 – Parameter Definition	40
Table 2 – Data cloning posterior estimates (1)	41
Table 3 – Data cloning posterior estimates (2)	42
Table 4 – Normalized data cloning posterior standard deviation	47

Contents

1	Introduction	9
2	Literature Review	11
2.1	DSGE Models	11
2.2	Estimation	13
2.3	Identification	16
3	Estimation and Identification Methodology: Data Cloning	21
4	The Smets and Wouters Model	24
4.1	The Model	24
4.2	Known Identification Issues	25
5	Data Cloning Application	27
5.1	Setup, Implementation and Results	27
5.2	Discussion	29
6	Conclusion	32
	Bibliography	33
	APPENDIX A Tables and Figures	39
	APPENDIX B The log-linearized Smets and Wouters (2007) model	51
B.1	The Sticky Price and Wage Equilibrium Conditions	51
B.2	The Flexible Price and Wage Equilibrium Conditions	55
B.3	The Exogenous Variables	56
B.4	The Steady-State Relations	56

1 Introduction

Over thirty years after Lucas (1976) Critique and the Real Business Cycle (RBC) of Kydland and Prescott (1982), and Dynamic Stochastic General Equilibrium (DSGE) models are widespread - central banks admittedly employ such models.¹ As Blanchard (2008) discuss, DSGE models are a product of the late twentieth century macroeconomics revolution, at its core lie the microfoundations of neoclassical growth theory, then augmented with market imperfections, real and nominal rigidities so that a better real-world characterization is achieved.

As discussed by Fernández-Villaverde (2010), DSGE models play a major role in modern macroeconomic research. Strong microfoundations make so that calibrated simulations can provide useful insights. The development of simulation-based MCMC estimation techniques powered by the Metropolis-Hastings (MH) algorithm, of novel optimization algorithms², and the steady decrease of the relative cost of computing hours contributed so that even more complex models (with over a hundred parameters like in de Castro et al. (2011)) can be estimated. As nothing is charge free, richer, more complex models come at costs. Canova and Sala (2009) motivate that due to the non-linear nature of DSGE systems, parameter identification issues often arise.

In the present paper we employ the data cloning algorithm developed by Lele et al. (2007) and Lele et al. (2012) to estimate the model of Smets and Wouters (2007) (SW). The data cloning algorithm allows the construction a numerical approximation of the maximum likelihood estimator (MLE) as the limit of Bayesian estimator based on MCMC simulations. To its implementation it is only necessary to build a vector containing K replications of the original samples, and correct the Fisher information matrix accordingly. As shown in Lele et al. (2007), under standard regularity conditions, with sufficiently large K , these estimators converge to the maximum likelihood estimators.

¹ Examples are the Federal Reserve Board (Erceg et al. (2005)), the European Central bank (Christoffel et al. (2007)), the Bank of England (Harrison et al. (2005)), and the Central Bank of Brazil (de Castro et al. (2011)).

² Such as the CSMINWEL of Sims (1999), and the Nelder and Mead (1965) simplex method.

The evaluation of the identification properties of the parameters of the model are a byproduct of the data cloning estimation method. Ill-identified parameters will display posterior volatility that does not decrease as the number of sample replicas increase. Such evaluation can be done by visually inspecting the data cloning posterior distribution of the parameters, or by calculating its normalized standard error. Also as discussed in Lele et al. (2007), the maximum eigenvalue of the data cloning posterior covariance matrix should converge to zero at the same rate as $1/K$, and provides a measure of global identification of the DSGE system.

Our results indicate that the SW model is, in general, not well identified. The posterior volatility of most parameters does not decrease as the number of clones increases. The posterior volatility of the data cloning parameter estimates is, in most cases, not reduced by an increase in the number of clones - thus few individual parameters can be said to be well-identified (namely ρ_m , ρ_b , σ_b , β , r_y , and γ). The maximum eigenvalue of the data cloning posterior covariance matrix does not converge to zero (and is consistently above its expected value of $1/K$), indicating that the SW model is not globally identified.

The data cloning algorithm we employ shares similarities with the bayesian learning rate indicator introduced by Koop et al. (2013). This procedure is used by Caglar et al. (2012) to study the identification properties of the SW model.

The rest of this paper is divided in five sections. In section 2 we overview the DSGE literature, emphasizing estimation techniques and identification problems. We describe the data cloning technique in Section 3. Section 4 contains a brief discussion of the SW model. Our results are reported and further discussed in section 5. Finally, we present some concluding remarks in section 6.

2 Literature Review

2.1 DSGE Models

DSGE models are a byproduct of the methodological revolution that started in the mid-1970s. The old Keynesian models are contested by Lucas (1976); on the grounds that because estimated parameters are not structural (in the sense that they derive from primitives of the model), they were not invariant given changes in specific economic policies, what makes old Keynesian models not suitable for policy analysis. On the empirical side, we see the popularization of multivariate autorregressive techniques as form to describe the data (think of Sims (1980)). The work of Kydland and Prescott (1982) is regarded as a particular cornerstone in the development of DSGE models because, as Fernández-Villaverde (2010) discusses, “for the first time, macroeconomists had a small and coherent dynamic model of the economy, built from first principles with optimizing agents, rational expectations, and market clearing, that could generate data that resembled observed variables to a remarkable degree”. Nowadays DSGE models are widely applied. They are the standard testing ground for modern macroeconomic theory, and Central Banks admittedly use them as tools to aid policy making¹.

As DeJong and Dave (2011) put it, DSGE models start com a characterization of the environment in which decision makers reside, a set of decision rules that dictate their behavior, and a characterization of the uncertainty they face when making decisions. This structure takes the form of a nonlinear system of expectational equations.

We will overview the DSGE literature with aid of a simple log-linearized example. Let $\hat{z}_t$ be a n_z -dimensional vector of stationary variables, and $\hat{z}^*$ be the steady-state values of $\hat{z}_t$. Following Iskrev et al. (2010), most linearized DSGE models can be represented as

$$\Gamma_0(\theta)z_t = \Gamma_1(\theta)\mathbb{E}_t z_{t+1} + \Gamma_2(\theta)z_{t+1} + \Gamma_3(\theta)\epsilon, \quad (2.1)$$

¹ Examples are the Federal Reserve Board (Erceg et al. (2005)), the European Central bank (Christoffel et al. (2007)), the Bank of England (Harrison et al. (2005)), and the Central Bank of Brazil (de Castro et al. (2011)).

where $z_t = \hat{z}_t - \hat{z}^*$, and the structural shocks ϵ_t are independent and identically distributed n_ϵ -dimensional random vectors with $\mathbb{E}\epsilon_t = 0$, $\mathbb{E}\epsilon_t\epsilon_t' = \mathbb{I}_{n_\epsilon}$. Matrices $\Gamma_0, \Gamma_1, \Gamma_2$ and Γ_3 are functions of a n_θ -dimensional vector of deep parameters θ , where θ is a point in the parameter space $\Theta \in \mathbb{R}^k$.

Since (2.1) depends on expectational components, it must be solved in terms of rational expectations. There are a number of algorithms for solving linear rational expectation models (see for example Blanchard and Kahn (1980), Anderson and Moore (1985), King and Watson (1998), Sims (2002)). Assuming that a unique solution exists, it can be written as

$$\underset{n_z \times 1}{z_t} = \underset{n_z \times n_z}{A(\theta)} \underset{n_z \times 1}{z_{t-1}} + \underset{n_z \times n_\epsilon}{B(\theta)} \underset{n_\epsilon \times 1}{\epsilon_t}, \quad (2.2)$$

where matrices A and B are unique for each value of θ .

Were all endogenous variables in z_t observable, The likelihood function of (2.2) could be directly recovered from data. Unfortunately the presence of unobserved variables is almost certain in practical applications (the capital stock is unobservable, for example). Thus the transition equation in (2.2) is augmented with the measurement equation (2.2), to yield the state-space representation of our DSGE model:

$$\underset{n_x \times 1}{x_t} = \underset{n_x \times n_z}{C(\theta)} \underset{n_z \times 1}{z_t} + \underset{n_x \times n_u}{D(\theta)} \underset{n_u \times 1}{u_t} + \underset{n_x \times 1}{\nu_t}, \quad (2.3)$$

where x_t is a n_x -dimensional vector of observed variables, u_t is a n_u -dimensional vector of exogenous variables, and ν_t is a n_x -dimensional random vector with $E\nu_t = 0$, and $E\nu_t\nu_t' = Q$, where $Q(\theta)$ is a $n_x \times n_x$ symmetric semi-positive definite matrix. And for a given value of θ , the matrices A , $\Omega \equiv BB'$, and $\hat{z}^*$ completely characterize the equilibrium dynamics and the steady-state properties of all endogenous variables in the linearized model.

Before estimation, it is important to ensure that observed variables must match their theoretical counterparts. This might involve the data's seasonal adjustment or trend removal, since observable variables must be stationary. But more crucially, measurement errors for the observable variables are included. Sometimes measurement errors are included to increase the number of structural shocks in the model, in order to avoid stochas-

tic singularity. Stochastic singularity happens when the number of observable variables is greater than the number of structural shocks observed in the model.

2.2 Estimation

Various methods for choosing parameter values for DSGE models have been proposed over the years. Early attempts traditionally involved informal calibration procedures (like in Kydland and Prescott (1982), Gregory and Smith (1991), Eichenbaum (1991)), that, as Kydland and Prescott (1996) define them, consist in matching a small number of selected real data and model features. Nevertheless a lot of effort has been, and is being, put in finding better, formal, ways of estimating DSGE models. The best distinction we are able to make here regards classical (frequentist) and Bayesian methods, but some procedures might be hybrid - like the data cloning we employ: here, Classical maximum likelihood estimates are obtained through Bayesian asymptotic theory.

The Generalized Method of Moments (GMM) is employed to estimate DSGE models in Hansen and Singleton (1983), Christiano and Eichenbaum (1992), Burnside et al. (1993). An alternative to GMM is the Simulated Method of Moments (SMM); proposed by McFadden (1989) and Pakes and Pollard (1989) to estimate discrete-choice problems and by Lee and Ingram (1991) and Duffie and Singleton (1990) to estimate time-series models, in such approach one chooses the parameter vector to minimize the distance between moments of actual and simulated data. Ruge-Murcia (2007) presents a good comparison of classical methods in estimating the one-sector real business cycle model; apart from GMM and SMM, the author also considers maximum likelihood and Simulated Quasi-Maximum Likelihood (SQML, or Indirect Inference).

First maximum likelihood estimations of DSGE model dates back to Hansen and Sargent (1980). Such approach consists in casting the DSGE model in state-space representation, like we did in equations and , and then reconstructing the likelihood function through filtering the data. Assume that the structural shocks ϵ_t are jointly Gaussian, then one can reconstruct the likelihood function of the data $X = [x_1, \dots, x_t]$ through Kalman

filter recursions. Let $e_{t|t-1} = x_t - C\hat{z}_{t|t-1} - Du_t$ be the one step-ahead forecast of the state vector $\hat{z}_{t|t-1}$, where $e_{t|t-1}$ is also Gaussian with zero mean and covariance matrix given by $S_{t|t-1} = CP_{t|t-1}C'$, where $P_{t|t-1} = \mathbb{E}(z_t - \hat{z}_{t|t-1})(z_t - \hat{z}_{t|t-1})'$ is the conditional covariance matrix of the one-step ahead forecast. It is now possible to write the log-likelihood function of data X as

$$\log(\mathcal{L}(X|\theta)) = \text{const.} - \frac{1}{2} \sum_{t=1}^T \log(\det(S_{t|t-1})) - \frac{1}{2} \sum_{t=1}^T e'_{t|t-1} S_{t|t-1}^{-1} e_{t|t-1}. \quad (2.4)$$

The Maximum Likelihood estimate of θ (MLE) is the value which maximizes (2.4)

$$\hat{\theta}^{ML} = \max_{\theta \in \Theta} \mathcal{L}(X|\theta). \quad (2.5)$$

The precision of $\hat{\theta}^{ML}$ is determined by the inverse of the Fisher information matrix, defined as

$$\mathcal{I}(X|\theta) := \mathbb{E} \left[\left(\frac{\partial \mathcal{L}(X|\theta)}{\partial \theta'} \right)' \left(\frac{\partial \mathcal{L}(X|\theta)}{\partial \theta'} \right) \right] \quad (2.6)$$

Unfortunately likelihood function of DSGE models are highly dimensional objects, with multiple local maxima and minima, and flat surfaces. Maximizing such function is a very difficult task, and pure maximum likelihood is not the standard estimation method of choice in DSGE models. As discussed by An and Schorfheide (2007), MLE are further subject to the “dilemma of absurd parameter estimate”: estimates of structural parameters generated with maximum likelihood procedures based only a set of observations are often at odds with additional information that the research may have.

Bayesian methods, powered by advances in computing, offer ways to estimate circumvent problems with ML estimation. As Fernández-Villaverde (2010) motivates, integrating a complicated function is actually easier than it is to maximize it, and is what is done in Bayesian exercises. Also, prior distributions of parameters are a formal way to introduce extra-sample information to the estimation process, possibly attenuating the “dilemma of absurd parameter estimate”.

Let $\pi(\theta)$ denote the joint prior distribution of the parameters θ . Remember that $\mathcal{L}(X|\theta)$ is the likelihood of the data X , given the parameters θ . Bayes theorem implies

that the posterior distribution of the parameters (given the data) is

$$\pi(\theta|X) = \frac{\mathcal{L}(X|\theta)\pi(\theta)}{\int_{\Theta} \mathcal{L}(X|\theta)\pi(\theta)d\theta}. \quad (2.7)$$

The denominator in (2.7) is integrated over all possible values of $\theta \in \Theta$, and thus constant. This value does not need to be evaluated in most Bayesian exercises, as we write the posterior as proportional to the likelihood times the prior (it is always possible to normalize (2.7) so that it is a proper distribution density function):

$$\pi(\theta|X) \propto \mathcal{L}(X|\theta)\pi(\theta). \quad (2.8)$$

As Fernández-Villaverde (2010) discusses, the growing popularity of Bayesian techniques is intertwined with developments in numerical simulation, specially Markov Chain Monte Carlo methods. In MCMC methods we are interested in producing a Markov chain whose ergodic distribution is $\pi(\theta|X)$. Once we have such chain we just have to simulate from it, and approximate $\pi(\theta|X)$ empirically.

The Metropolis-Hasting (MH) algorithm provides a clever way to construct such chain. We adopt a proposal distribution and an initial value for the parameter vector, then draw a new parameter vector from the proposal distribution. If the new parameter value is more likely than the previous one, then surely the parameter vector is updated; if the new parameter is less likely, then sometimes the algorithm will update the parameter and sometime not. This means that the chain will always move up to regions of higher probability, but will also sometimes move down to regions of lower probability. Under certain regularity conditions, the chain created by this method will have $\pi(\theta|X)$ as ergodic distribution. Here we adapt the pseudo-code presented in Fernández-Villaverde (2010) for a simple MH algorithm:

- Step 0. Initialization: Set $i = 0$ and an initial parameter vector θ_i . Solve the model for θ_i , build the state space representation (equations (2.2) and (2.2)). Evaluate the prior distribution $\pi((\theta)_i)$ and the likelihood function $\mathcal{L}(X|\theta_i)$. Set $i = i + 1$.
- Step 1, Proposal Draw: Get a draw θ_i^* from a proposal distribution density function $q(\theta_{i-1}, \theta_i^*)$

- Step 2, Solving the Model: Solve the model for θ_i^* and build the new state space representation.
- Step 3, Evaluating the proposal: Evaluate $\pi(\theta_i^*)$ and $\mathcal{L}(X|\theta_i)$ with the Kalman filter (like in (2.4)).
- Step 4, Accept/Reject: Draw $\Xi_i \sim U(0, 1)$. If $\chi_i \leq \frac{\mathcal{L}(X|\theta)\pi(\theta)q(\theta_{i-1}, \theta_i^*)}{\mathcal{L}(X|\theta^*)\pi(\theta^*)q(\theta_i^*, \theta_{i-1})}$ set $\theta_i = \theta_i^*$, otherwise set $\theta_i = \theta_{i-1}$.
- Step 5, Iteration: Repeat steps 1-4 until a good characterization of $\pi(\theta|X)$ is obtained.

2.3 Identification

Identification is defined by Canova and Sala (2009) as having “to do with the ability to draw inference about the parameter of a theoretical model from an observed sample”. This vague definition is congruent with the elusive nature of identification question in DSGE models, as it seldom has a yes-or-no answer. The classical identification results of Rothenberg (1971), where local identifiability of a parameter vector is associated with non-singularity of the Fisher information matrix, do not apply to DSGE models. As shown by, Komunjer and Ng (2011), because the reduced form parameters of the state-space representation are not generally identifiable, the non-singularity of the information matrix does not suffice to ensure local identification of the parameters.

Problems might occur at all different stages of the implementation of a DSGE model described above. Canova and Sala (2009) introduce four operators that interact in the process of mapping the data X to the estimated parameter vector $\hat{\theta}_T^{ML}$. These are not mathematical operators, but rather a description of specific stages in the implementation of a DSGE experiment. We think those help understand the specific identification problems that we will cover:

- 1. The “solution” mapping, linking the parameters θ to the coefficients of the solution. In our example it is represented by matrices A and B . This mapping could be

ill-behaved because structural parameters disappear from the solution, do not have independent variation, or induce small changes in the coefficients of the solution.

- 2. The “moment” mapping, linking the coefficients of the solution to the functions of interest (impulse responses in our case). The problem here could be that the selection of particular impulse responses poorly packages the information contained in the solution.
- 3. The “objective function” mapping, linking the functions of interest to the population objective function. This mapping could be ill-behaved when the objective function display multiple maxima or is insensitive to changes in values of specific parameters.
- 4. The “data” mapping, linking the population and the sample objective functions. Here it could be that estimated responses are inconsistent estimators of the population ones—and this may occur, for example, because of errors in the identification of shocks—or when a subset of the endogenous variables appearing in the solution is omitted from the estimated VAR.

Problems arise when some parameters disappear from the objective function, or enter it only proportionally.

Definition 1. Let $\theta = [\theta_1, \theta_2]$ and partition $\Theta = [\Theta^1, \Theta^2]$. The parameters θ_1 are said to be **locally under-identified** around some neighborhood Θ_1 if $\mathcal{L}(X, m_1, \theta_1, \theta_2) = \mathcal{L}(X, m_2, \theta^2)$ for all $\theta_1 \in \Theta_1 \subset \Theta_2$.

Definition 2. Let $\theta = [\theta_1, \theta_2]$ and partition $\Theta = [\Theta^1, \Theta^2]$. The parameters θ_1 are said to be **locally partially identified** if $\mathcal{L}(X, m_1, \theta) = \mathcal{L}(X, m_2, \theta_1 f(\theta_2))$ for all $\theta_1 \in \Theta_1 \subset \Theta^1$, and all $\theta_2 \in \Theta_2 \subset \Theta^2$.

It is not uncommon for approximation methods to induce under- or partial identification problems. For instance: SW model that we employ in our exercise has two pairs of parameters that, due to log-linear approximation, are only partially identified. One can

see from Equations (B.15) and (B.18) in Appendix B that parameters ξ_p and ϵ_p , ξ_w and ϵ_w enter the model solution only proportionally².

Partial and under-identification are problems of the model, the objective function, and/or the approximation method employed; therefore a model which display this sort of problems, will display them no matter the data-set used in the application. Or we could say that they involve only the "solution" and the "objective function" mappings of Canova and Sala (2009).

One fundamental identification problem arises when the likelihood function does not have an unique maximum. Thus under the point of view of the objective function, two models, potentially having different economic interpretations are indistinguishable, and informational external to the model needs to be introduced in order to choose between structural forms.

Definition 3. *Two models m_1 and m_2 , with parameter vectors θ^1 and θ^2 , are **observationally equivalent** given data if $\mathcal{L}(X, m_1, \theta^1) = \mathcal{L}(X, m_2, \theta^2)$.*

Observational equivalence is a well known phenomenon in macroeconomics, and has played part in many debates over the years. For example Sargent (1976) showed that it is difficult to distinguish between old Keynesian and neoclassical models using only parameters estimated from a single policy regime. Chari et al. (2008) uses observational equivalent, but structurally different, models to question to what extent shocks in the SW model are interpretable.

Regarding the fourth operator of Canova and Sala (2009), the “data” mapping. Problems here might occur because of errors in the identification of shocks, or when the set of observable variables is not informative about population characteristics. The SW model has problems with the identification of the standard error of wage and price mark-up innovations that has lead to criticism, we explore this discussion in Sections 4 and 5. Canova et al. (2014) and discuss two methods for choosing the observable variables to estimate singular DSGE models.

² The parameters ϵ_p and ϵ_w are calibrated before estimation. We discuss details from the estimation in Section 5.

The last specific problem of identification we are going to overview differs from the previous three in that it does not have a straight yes-or-no answer. And could be thought of as a defect on the curvature of the objective function (Canova and Sala (2009)).

Definition 4. *A parameter vector θ is locally weakly identifiable if there exists a unique θ^* such that $\mathcal{L}(X, \theta^*) = 0$ but $\mathcal{L}(X, \theta^*) < \epsilon$ for all $\theta^0 \in \Theta_1 \subset \Theta$ and all $\epsilon > 0$. And globally weakly identifiable if this occur for all $\theta \in \Theta$.*

Iskrev (2010) approaches the weak identification problem as a model feature, both in mathematical and economic sense. He discusses that although the non-singularity of $\mathcal{I}(X|\theta)$ ensures that the expected likelihood is not flat and achieves a locally unique maximum; the precision with which the true parameter θ may be estimated in finite samples depends on the degree of curvature of the likelihood around an open neighborhood of $\hat{\theta}$, to which the rank condition provides no information. If the curvature of the log-likelihood is nearly flat around $\hat{\theta}$, then large changes in $\hat{\theta}$ will induce small changes in $\mathcal{L}(X, m_1, \hat{\theta})$, due to random variations in the sample, which result in relatively large changes in the value of θ that maximizes the observed likelihood function. The economic argument the author presents is that a weakly identified parameter is one which is either irrelevant, because it has only a negligible effect on the likelihood, or nearly redundant, because its effect on the likelihood may be closely approximated by other parameters - and thus the value of such parameters is hard to pin down on the basis of the information contained in the likelihood function.

Although the Bayesian approach has compelling advantages over the Frequentist when estimating DSGE models, the former introduces new elements that could complicate inference. In the Bayesian exercise we are interested in the posterior $\pi(\theta|X)$, which is (proportional to) the likelihood $\mathcal{L}(X|\theta)$ times the prior $\pi(\theta)$. The posterior is defined as long as there is a proper prior; i.e., the limiting case is inference without data, in which case the posterior is equal to the prior. Thus, as An and Schorfheide (2007) discuss, priors “might add curvature to a likelihood function that is (nearly) flat in some dimensions of the parameter space and therefore strongly influence the shape of the posterior distribution”.

A naive Bayesian identification test is visually contrasting prior and posterior shape. Intuitively, if the posterior is very similar to the prior, then data had little to say about the parameter of interest. Canova and Sala (2009) discuss that this is only true if the parameter space is variation free; i.e., there are no model-implied restrictions on the parameter space. DSGE models seldom are variation free, they usually invariably impose stability restrictions or economic motivated non-negativity constraints on some parameters. If the parameter space is not variation free, then belief updates about one parameters might affect the parameter space of another, ultimately affecting the shape of the posterior distribution - even if data was not informative.

Parameter identification problems can also be investigated via computational methods. Koop et al. (2013) introduce two Bayesian identification indicators. The first, the Bayesian comparison indicator looks to the expected posterior shape relative from prior and provides a yes-or-no answer to the identification problem. The second one, the Bayesian learning rate indicator, consists of augmenting the observed sample with model simulated paths, and then observing the standard error of MCMC estimates as the simulated sample increases. The second procedure proposed by Koop et al. (2013) has many similarities with the data cloning method described in the next section.

3 Estimation and Identification Methodology: Data Cloning

The Data Cloning (DC) methodology was developed by Lele et al. (2007) and Lele et al. (2012), as a computational method to obtain ML estimators as the limit of Bayesian estimators when the original data is pooled many times. It was first designed to estimate complex ecological systems, but then applied to estimate a wide range of Generalized Linear Mixed Models. Furlani et al. (2013) demonstrate the possibility of estimating DSGE models through the DC method.

Lele et al. (2012) expose the data-cloning algorithm with aid of a thought experiment: imagine a hypothetical situation where the statistical experiment underlying the observations $X_{(n)}$ is repeated independently K times, and coincidentally all realizations were $X_{(n)}$. The likelihood function based on the combination of the data from these K independent experiments is given by $\mathcal{L}(X|\theta)^K$. We can modify (2.7) such that the posterior function obtained through the data-cloning method is

$$\pi(\theta|X) = \frac{\mathcal{L}(X|\theta)^K \pi(\theta)}{\int_{\Theta} \mathcal{L}(X|\theta)^K \pi(\theta) d\theta}. \quad (3.1)$$

Lele et al. (2012) prove that, following from the standard result regarding the asymptotic behavior of the posterior distributions (e.g., Walker 1969), under regularity conditions, if K is large, then

$$\pi_K(\theta|X) \xrightarrow{K \rightarrow \infty} \mathcal{N}\left(\theta, \frac{1}{K} \mathcal{I}(\theta)^{-1}\right). \quad (3.2)$$

This distribution degenerates at the MLE θ^{ML} , defined in (2.5), also the mean of the DC posterior $\pi_K(\theta|X)$ is the MLE and K times the posterior variance is the corresponding variance of θ^{ML} .

Sampling from $\pi_K(\theta|X)$ is easy using the MH algorithm, and it requires no additional coding, we simply conduct the thought experiment described above using a computer. We create K -cloned datasets, $X^K = (X, \dots, X)$, by repeating the observed data

vector K times. Then pretend that these data were obtained from K independent experiments and use the standard MCMC method to sample from the posterior distribution $\pi_K(\theta|X)$. Through this indirect procedure, one is capable of obtaining the maximum likelihood estimator while also avoiding: having to evaluate high-dimensional integrals, maximize the likelihood function, or compute the curvature of the likelihood function.

It is possible to evaluate the individual identification of estimated parameters directly through the data-cloning procedure. As discussed in Lele et al. (2007) and Lele et al. (2012), a parameter is nonidentifiable if its posterior variability does not converge to zero as the number of clones increase. This can easily be checked by observing the standard errors of the chains sample in the MCMC, and how the standard error diminishes (or not) as K gets bigger; or by visually inspecting the posterior distribution of each parameter - that if identified, should have a posterior distribution that degenerates around the ML estimate.

Let λ_K denote largest eigenvalue of the posterior variance of $\pi_K(\theta|X)$. Lele et al. (2012) shows that λ_K converges to zero as the same rate as $1/K$. If we divide λ_K by λ_1 (the largest eigenvalue of the single-sample posterior variance), we have what Lele et al. (2012) calls the standardized largest eigenvalue λ_K^S . It is then possible to compare the value of λ_K^S with $1/K$ - the expected value for λ_K^S - in order to obtain a global measure of identification for the DSGE system.

Artificially increasing sample size by repeating observed outcomes of macroeconomic variables does not make sense. These variables display temporal dependence, and are subject to structural breaks of all sorts. In order to set ideas straight, and avoid such criticism, we repeat Lele et al. (2012) and emphasize that the data cloning is nothing but a computational algorithm to compute $\hat{\theta}^{ML}$ and $\mathcal{I}(\theta)$. As shown in Lele et al. (2007), although the proposed thought experiment alludes to it, the mathematical proof of the algorithm does not depend and in no way assumes that the K clones are independent from each other. As the number of clones increases, the algorithm better approximates the true location of the likelihood function and the true inverse of the Fisher information for the observed data. The statistical accuracy of the estimator is a function of sample

size and is not enhanced by the data cloning procedure, increasing the number of clones will only contribute to enhance the numerical approximation of the ML estimates.

A compelling theoretical feature of the data cloning method is to eliminate prior influence on posterior estimates. As discussed in section 2, prior distributions are sort of two-edged sword: on one side they provide a formal way to introduce extra-sample information into the regression; but even diffuse priors induce curvature into the likelihood function, potentially clouding parameter identification problems.

Furlani et al. (2013) discuss that the quasi-Bayesian Laplace type estimators presented by Chernozhukov and Hong (2003) relates to the data-cloning method, and why the former displays more desirable properties. The LTE is equivalent to the maximum likelihood estimator when using weighting functions equivalent to improper priors; in the case of proper priors (weights), the LTE is only asymptotically equivalent to the maximum likelihood estimator, but as they are based on the original data sample without replications, they still carry the information of the pseudo-priors used.

4 The Smets and Wouters Model

4.1 The Model

Following an apparent consensus in recent literature, we take the model presented by SW as a good representative of the class of New-Keynesian (NK) DSGE models, and a proper testing ground for the Data Cloning methodology. The model indeed contains a plethora of real and nominal rigidities that are emblematic of NK-DSGE models. Firms and households have some monopoly power over wages and prices; capital utilization is variable, but there growing are costs to changing it; not all prices and wages are reoptimized each period, and stagnant prices are partially indexed to past inflation. Here we merely describe the model basic interactions, variables and parameters, present the log-linearized equations, and discuss some known identification issues with the SW model. A more detailed exposition of the model is present in Annex B, and a complete derivation can be found in the on line appendix of the original paper.

The SW economy is composed of many households and firms that differentiate labor and goods they supply, and a Taylor-Like monetary authority. Infinitely-lived households maximize an utility function unseparable on both arguments, consumption and labor effort. Utility derived is derived from current consumption relative to past aggregate consumption, characterizing habit formation. Labor is differentiated between households, so that an explicit wage equation for real wages exist. Households choose how much capital to accumulate given the portfolio adjustment costs they face, and rent capital services to firms. Firms determine labor and capital inputs to produce differentiated goods. Capital stock utilization is variable and changes subject to growing capital adjustment costs. Dynamics of wages and prices are enriched by an indexed-Calvo mechanism. It is, only a subsample of households and firms reoptimize their prices or wages each period, and non-reoptimized prices and wages are partially indexed to past inflation. The monetary authority sets the nominal interest rate while concerned with the past interest rate, cur-

rent inflation rate, the output gap, and the difference between current and past output gaps.

The model contains 7 are observable variables: real output, worked hours, inflation rate, consumption, investment, and capital stock (a better characterization of the data and the measurement equations can be found in Annex B). Seven exogenous structural shocks match the observable variables and are described in Equations (B.21a) to (B.21g). There are 41 parameters in the SW model: 17 dictate the exogenous shock processes, and the remainder characterize the deep structural features of the economy. Of those, 5 are calibrated, leaving a total of 36 parameters to be estimated. Table 1 contains descriptions of the parameters, and the log-linearized equations are exposed in Annex B.

4.2 Known Identification Issues

Iskrev (2010) and Komunjer and Ng (2011) analyze the identification of the SW model through rank tests on the information matrix and the spectral density matrix. They reach the same conclusion, that two pairs of parameters are partially identified (ξ_w and ϵ_p , ξ_w and ϵ_w). This is the reason why Smets and Wouters (2003) calibrate both Kimball (1995) curvature parameters. The steady-state labor market mark-up, ϕ_w is identified at the prior and posterior means, although Iskrev (2010) finds 3 points (out of more than 900 thousand random draws) from the parameter space where it is not.

Iskrev et al. (2010) study the identification strength of the SW model by evaluating the expected curvature of the log-likelihood function. He finds that at the posterior mean, the likelihood function is little sensitive to the parameters $\bar{l}, \beta, \rho_r, \sigma_l, \iota_p, r_t$, and δ ; also, pairwise correlation analysis shows that ϕ_w had a 0.94 correlation with ξ_w . On the other hand, the best identified parameters, according to the same measure are $\rho_g, \rho_w, \rho_a, \bar{\gamma}, \rho, \rho_p, \sigma_g, \mu_w$, and σ_r , that is, the deterministic growth trend and the parameters related to the dynamics of exogenous shocks.

Caglar et al. (2012) apply the Bayesian learning rate indicator presented by Koop et al. (2013) to computationally evaluate the identification strength of the SW model.

Their method is close to the data-cloning algorithm, and so are their conclusions to ours. They find that only 8 of the 36 parameters can be said to be identified, namely: $\bar{\gamma}$, ρ_g , ρ_a , ρ_w , μ_w , ρ_p , σ_l , σ_w . The worst identified parameters are $\bar{\pi}$, r_π , $r_{\Delta y}$, $\bar{l}$, β , r_y , σ_l , and ψ .

A different line of criticism over the SW model regards the implausibly high volatility that the model attributes to wage mark-up shocks (ε_t^w). Chari et al. (2009) notice that the wage mark-up shock dominates long-run output and employment fluctuations. The authors then show a model observational equivalent to SW's, but where the mark-up shocks have different economic implications. Chari et al. (2009) argue that because the SW model can not distinguish between wage mark-up and labor supply shocks, it is not suitable for policy analysis. Galí et al. (2011) respond to the critique of Chari et al. (2009) by reformulating the SW model so that the unemployment rate is included as an observable variable. In the reformulated model, Galí et al. (2011) find a diminished role for wage mark-up shocks as a source of output and employment fluctuations, even though those shocks preserve a large role as drivers of inflation.

Evidences from microdata indicate that the indexed-Calvo scheme, that characterize price and wage setting in the SW model, do not well-characterize firm level behavior (Bils and Klenow (2004), Dixon and Kara (2010)). Bils et al. (2012) argue that this specification damp inflation volatility, resulting in high price mark-up shocks (ε_t^p) required to explain observed data. Kara (2015) estimates a modified version of the SW model in which partial indexation is removed from the Calvo pricing structure, but there are different reoptimization probabilities among sectors of the economy. To do such he uses the dataset of Bils et al. (2012), which is based on U.S. Consumer Price Index microdata, and calibrates reoptimization probabilities for 10 sectors. Kara (2015) finds that with such specification, ε_t^p reduces significantly, from 0.91 to 0.24.

5 Data Cloning Application

5.1 Setup, Implementation and Results

The present application of the data cloning method consists in first replicating the baseline exercise of SW, and then estimating the model using gradually more sample clones - we consider 2, 3, 5, 10, and 25. Local identification will be assessed by observing the posterior volatility of the data cloning-estimated parameters. We also measure the global identification identification properties of the DSGE system obtained through the maximum eigenvalue of the data cloning posterior covariance function.

Calculations were carried out in MatLab, with the help of Dynare. The implementation of the algorithm requires no further coding, as it consists simply in chaining the observed data set an arbitrary number of times. That said the algorithm is computationally expensive, since the time required for the MCMC sampling of the posterior distribution increases considerably with the number of clones. Also, ensuring convergence of the Markov Chain here is a tricky part nevertheless. The Metropolis-Hastings algorithm requires the researcher to set a scaling constant before running. This constant, *MH_JSCALE* in Dynare's estimation built-in function, has a significant effect on the acceptance ratio of the MH algorithm. There is consensus in the literature that a stable acceptance ratio of about one-third is desirable in order to ensure a well-mixed chain. The problem here is that reasonable values for this constant vary with the number of sample replicas employed, having to be calibrated through trial and error. For example, in the single-sample baseline exercise the default 0.3 value of *MH_JSCALE* works fine, but with 25 clones we needed to set a scaling constant of 0.0005. Further detail and the codes for reproduction are provided in the supplement material.

Tables 2 and 3 display our estimation result. The first three columns regard prior specification, and the fourth column contains the results from numerical optimization of the posterior function. Most of the remaining columns report the posterior mean and

standard errors for the single-sample and data cloning estimations. To obtain the latter statistic we multiply the standard error obtained from the MCMC simulation with K clones by $\sqrt{K}$. For quick reference, the benchmark values (the ones that SW obtained) can be found in the last column of Tables 2 and 3. Data cloning point estimates are mostly in line with the benchmark values. The only parameters whose data cloning estimated mean fall outside the 95% confidence interval provided by SW are σ_l , α , and to some extent μ_p . None of these parameters fare well in our identification test, so exploring the economic implications of shifts in their values would be inappropriate. In fact, our results indicate that identification problems are widespread, extending to both shock-related and deep parameters.

We report the posterior distribution plots for the single sample estimate in Figures 1 to 4, and for the 25-clones estimate in Figures 5 to 8. As discussed in Section 3 a visual inspection of the posterior distribution is a valid identification test in the data cloning context. It seems that 25 replicas is enough for some parameters posterior distribution to nearly degenerate - in the sense that they can be ill-behaved (and in fact many display local maximum), but the whole probability mass is concentrated in a small interval around the posterior mean (hopefully, the location of the likelihood function). Parameters with clear non-collapsing posteriors are ψ , h , ξ_w , σ_l , ξ_p , Ψ , ϕ , r_π , $r_{\Delta y}$, $\hat{\pi}$, $\hat{l}$, α . Figures 1 and 2 show that most shock-related parameters have degenerate posterior - but, as we will see, some of them are problematic nevertheless.

The visual procedure should be equivalent to observing the posterior volatility as the number of clones increases. So following Furlani et al. (2013), we calculate normalized standard errors (proportional to the standard error in the single-sample estimation) to evaluate whether the data cloning algorithm is successful in reducing posterior volatility. Table 4 displays the normalized standard errors as a function of the number of clones, denoted by s_K^* . To enhance visualization we list the values in decreasing order, relative to the 25-clones estimate. Thus Ψ is the parameter with the greatest normalized standard error after 25 clones, and σ_b is the one with smallest - but this does not mean that the parameters in the bottom of list are well identified. We plot these values in Figures 9 to

12, and from the graphs it becomes clear that for no parameter we have that an increase in the number of clones always reduced the normalized standard deviation. We take this erratic behavior of posterior volatility as an indication of bad model identification.

In Figure 13 we plot the standardized maximum eigenvalue of the data cloning posterior covariance matrix λ_K^S against its expected value of $1/K$. As discussed in Section 3, such comparison provides a measure of global identification of the model. We find that λ_K^S is above its reference level for every number of clones employed. Furthermore, the maximum eigenvalue does not seem to be converging to zero - since λ_K^S actually increases from 3 to 5 clones, and appears to reach a lower bound at around 0.15. From these evidence we can not say that the SW model is globally identified, what is to be expected, given the poor individual parameter identification performance.

5.2 Discussion

Our results are disappointing, in the sense that the vast majority of the parameters in the SW model cannot be demeaned identified under the standards of our exercise. We can say that the structure of exogenous monetary policy (ρ_m and σ_m), and for preference-shifting (ρ_b , σ_b) processes is identified, and that structural parameters β , r_y , and γ are relatively well-identified. Remember that these parameters are well identified in the sense that their data cloning posterior distribution visually collapses round the posterior mean, and their normalized standard error diminishes as the number of clones increased. Since the model seems to me ill-identified as a whole it is not sure if we should even trust these positive results. In fact Iskrev et al. (2010), when evaluating the identification strength of SW, a found that the subjective discount rate β has little expected impact over the likelihood function, and should be specially difficult to identify. So we will relate our results with the literature and try to perhaps distinguish between different sources of ill-identification.

Although it concentrates more identified parameters than the structural part of the model, some parameters of the shock structure are problematic nevertheless. ARMA

parameters ρ_a , ρ_g , ρ_π , ρ_w , and μ_p have data cloning posterior distributions that degenerate close to its upper bound of 1, potentially indicating non-stationarity in the exogenous shock processes. Visually these parameters seem identified, but their standard errors display erratic behavior. Also, it is outside the scope of the present exercise to discuss what it means for an autorregressive parameter to be identified at one, as it has potentially different implications for Bayesian and classical statistics (see Sims (1988) for a discussion). We could speculate, however, that this happens because the simple AR structure is unable to account for the dynamics of non-specific technology and government spending. Also, as discussed in Section 4, Bils et al. (2012) and Chari et al. (2009) argue that the standard deviation of mark-up innovations σ_p and σ_w are not identified; we do not find evidences to contradict them, since although these parameters are visually identified, their posterior volatility is not reduced. This could be the source of non-stationarity in the price and wage mark-up processes.

The parameters that determine price and wage rigidities, i.e., the Calvo reoptimization probabilities ξ_p and ξ_w , and the degree of partial indexation ι_p and ι_w , are not well identified. This could be due to high collinearity between them, as suggest Iskrev (2010), Canova and Sala (2009). But it is also important to remember that price rigidities are not structural model features (in the sense that they don't derive from primitives), and there is nothing that justifies the introduction of these rigidities besides a higher level of realism in the model. However, the empirical validity of the indexed-Calvo model is questionable (Dixon and Kara (2010), Bils et al. (2012), Bils and Klenow (2004)), so it could be that this part of the model is badly specified. Caglar et al. (2012) studied the identification of the SW model through an exercise that is similar to ours, instead of replicating the observed dataset they apply the method of Koop et al. (2013) and employ model-simulated series to augment the sample. They find that the investment adjustment cost ψ and the elasticity of labor supply σ_l are highly collinear pairwise, and Motula finds similar results in a SMM exercise. Two out of the four trend parameters are only weakly identified: the steady-state inflation rate $\hat{\pi}$, and the steady-state worked hours $\hat{l}$. As discussed in Caglar et al. (2012) and Canova and Sala (2009), trends level parameters

mainly affect the first moments, and having constant terms change the identification in general. Caglar et al. (2012) conjecture that, if we do not subtract the means from the log-linearized variables and instead add constant terms in the equations, the identification of these parameters could improve significantly.

The monetary policy reaction rule coefficients regarding past output gap $r_{\Delta y}$, and past inflation r_{π} display non-collapsing priors in Figures 7 and 8, and are thus regarded as ill-identified. This is very undesirable, since r_{π} is the most important parameter in the Taylor-rule. In their exercise, Caglar et al. (2012) also find Taylor rule parameters to be weakly identified. The authors argue that this is not surprising, that the literature has extensively discussed how a simplified Taylor rule that reacts only to inflation performs as well as a full rule, such the one of SW; and that it is hard to discern from data whether the monetary authority responds to changes inflation or output.

6 Conclusion

We employed the data cloning methodology of Lele et al. (2007) to estimate the Smets and Wouters (2007) model, and then study its identification properties. We find point estimates mostly in line with the original results, but the SW model appears to be only poorly identified. The posterior volatility of the data cloning parameter estimates is, in most cases, not reduced by an increase in the number of clones - thus few individual parameters can be said to be well-identified (namely ρ_m , ρ_b , σ_b , β , r_y , and γ). Not surprisingly, the maximum eigenvalue of the data cloning posterior covariance matrix does not converge to zero (and is consistently above its expected value of $1/K$), indicating that the SW model is not globally identified.

Possible causes for the identification problems were discussed. Price and wage rigidity parameters ξ_p , ξ_w , ι_p and ι_w , are collinear (Iskrev (2010), Canova and Sala (2009)). Also, microevidence suggest the indexed-Calvo structure poorly reflects firm-level behavior (Dixon and Kara (2010), Bils et al. (2012), Bils and Klenow (2004)). Some trend parameters ($\hat{\pi}$ $\hat{l}$) are problematic, but trends parameters mainly affect the first moments, and having constant terms change the identification in general (Caglar et al. (2012), Canova and Sala (2009)). The investment adjustment cost ψ and the elasticity of labor supply σ_l are collinear pairwise (Caglar et al. (2012)). Mark-up shocks σ_p and σ_w are disproportional to reality and economically non-interpretable (Bils et al. (2012), Chari et al. (2009)). Some parameters of the ARMA structure have estimates close to their upper bound of 1, namely ρ_a , ρ_g , ρ_π , ρ_w , and μ_p . The coefficients of the monetary policy reaction function are only poorly identified, specially $r_{\Delta y}$, r_π ; it is hard to tell from data whether the monetary authority responds to changes inflation or output (Caglar et al. (2012)).

Bibliography

- AN, S.; SCHORFHEIDE, F. Bayesian analysis of dsge models. *Econometric reviews*, v. 26, n. 2-4, p. 113–172, 2007.
- ANDERSON, G.; MOORE, G. A linear algebraic procedure for solving linear perfect foresight models. *Economics letters*, Elsevier, v. 17, n. 3, p. 247–252, 1985.
- BILS, M.; KLENOW, P. J. Some evidence on the importance of sticky prices. *Journal of Political Economy*, v. 112, n. 5, 2004.
- BILS, M.; KLENOW, P. J.; MALIN, B. A. Reset price inflation and the impact of monetary policy shocks. *The American Economic Review*, American Economic Association, v. 102, n. 6, p. 2798–2825, 2012.
- BLANCHARD, O. J. *The state of macro*. [S.l.], 2008.
- BLANCHARD, O. J.; KAHN, C. M. The solution of linear difference models under rational expectations. *Econometrica: Journal of the Econometric Society*, p. 1305–1311, 1980.
- BURNSIDE, C.; EICHENBAUM, M.; REBELO, S. Labor hoarding and the business cycle. *Journal of Political Economy*, v. 101, n. 2, p. 245–273, 1993.
- CAGLAR, E.; CHADHA, J. S.; SHIBAYAMA, K. Bayesian estimation of dsge models: Is the workhorse model identified? In: KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM WORKING PAPER SERIES. [S.l.], 2012.
- CALVO, G. A. Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, Elsevier, v. 12, n. 3, p. 383–398, 1983.
- CANOVA, F.; FERRONI, F.; MATTHES, C. Choosing the variables to estimate singular dsge models. *Journal of Applied Econometrics*, Wiley Online Library, v. 29, n. 7, p. 1099–1117, 2014.
- CANOVA, F.; SALA, L. Back to square one: identification issues in dsge models. *Journal of Monetary Economics*, Elsevier, v. 56, n. 4, p. 431–449, 2009.
- CASTRO, M. R. de et al. *SAMBA: Stochastic analytical model with a bayesian approach*. [S.l.], 2011.
- CHARI, V.; KEHOE, P.; MCGRATTAN, E. New keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, v. 1, n. 1, p. 242–266, 2009.
- CHARI, V. V.; KEHOE, P. J.; MCGRATTAN, E. R. *New Keynesian models: not yet useful for policy analysis*. [S.l.], 2008.
- CHERNOZHUKOV, V.; HONG, H. An mcmc approach to classical estimation. *Journal of Econometrics*, Elsevier, v. 115, n. 2, p. 293–346, 2003.

- CHRISTIANO, L. J.; EICHENBAUM, M. Current real-business-cycle theories and aggregate labor-market fluctuations. *The American Economic Review*, p. 430–450, 1992.
- CHRISTOFFEL, K.; COENEN, G.; WARNE, A. Conditional versus unconditional forecasting with the new area-wide model of the euro area. In: *Central Bank Workshop on Macroeconomic Modelling*. [S.l.: s.n.], 2007.
- DEJONG, D. N.; DAVE, C. *Structural macroeconometrics*. [S.l.]: Princeton University Press, 2011.
- DIXON, H.; KARA, E. Can we explain inflation persistence in a way that is consistent with the microevidence on nominal rigidity? *Journal of Money, Credit and Banking*, Wiley Online Library, v. 42, n. 1, p. 151–170, 2010.
- DUFFIE, D.; SINGLETON, K. J. *Simulated moments estimation of Markov models of asset prices*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1990.
- EICHENBAUM, M. Real business-cycle theory: wisdom or whimsy? *Journal of Economic Dynamics and Control*, Elsevier, v. 15, n. 4, p. 607–626, 1991.
- ERCEG, C. J.; GUERRIERI, L.; GUST, C. J. Sigma: a new open economy model for policy analysis. *Board of Governors of the Federal Reserve System Research Paper Series*, 2005.
- FERNÁNDEZ-VILLAYERDE, J. The econometrics of dsge models. *SERIEs*, Springer, v. 1, n. 1-2, p. 3–49, 2010.
- FURLANI, L. G. C.; LAURINI, M. P.; PORTUGAL, M. S. Data cloning: Maximum likelihood estimation of dsge models. In: *35th Meeting of the Brazilian Econometric Society*. [S.l.: s.n.], 2013.
- GALÍ, J.; SMETS, F.; WOUTERS, R. *Unemployment in an estimated new keynesian model*. [S.l.], 2011.
- GREGORY, A. W.; SMITH, G. W. Calibration as testing: Inference in simulated macroeconomic models. *Journal of Business & Economic Statistics*, v. 9, n. 3, p. 297–303, 1991.
- HANSEN, L. P.; SARGENT, T. J. Formulating and estimating dynamic linear rational expectations models. *Journal of Economic Dynamics and Control*, Elsevier, v. 2, p. 7–46, 1980.
- HANSEN, L. P.; SINGLETON, K. J. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *The Journal of Political Economy*, p. 249–265, 1983.
- HARRISON, R. et al. *The Bank of England quarterly model*. [S.l.]: Bank of England London, 2005.
- ISKREV, N. Local identification in dsge models. *Journal of Monetary Economics*, Elsevier, v. 57, n. 2, p. 189–202, 2010.
- ISKREV, N. et al. *Evaluating the strength of identification in DSGE models. An a priori approach*. [S.l.], 2010.

- KARA, E. The reset inflation puzzle and the heterogeneity in price stickiness. *Journal of Monetary Economics*, Elsevier, v. 76, p. 29–37, 2015.
- KIMBALL, M. S. *The quantitative analytics of the basic neomonetarist model*. [S.l.], 1995.
- KING, R. G.; WATSON, M. W. The solution of singular linear difference systems under rational expectations. *International Economic Review*, p. 1015–1026, 1998.
- KOMUNJER, I.; NG, S. Dynamic identification of dynamic stochastic general equilibrium models. *Econometrica*, p. 1995–2032, 2011.
- KOOP, G.; PESARAN, M. H.; SMITH, R. P. On identification of bayesian dsge models. *Journal of Business & Economic Statistics*, v. 31, n. 3, p. 300–314, 2013.
- KYDLAND, F. E.; PRESCOTT, E. C. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, p. 1345–1370, 1982.
- KYDLAND, F. E.; PRESCOTT, E. C. The computational experiment: an econometric tool. *The Journal of Economic Perspectives*, p. 69–85, 1996.
- LEE, B.-S.; INGRAM, B. F. Simulation estimation of time-series models. *Journal of Econometrics*, Elsevier, v. 47, n. 2, p. 197–205, 1991.
- LELE, S. R.; DENNIS, B.; LUTSCHER, F. Data cloning: easy maximum likelihood estimation for complex ecological models using bayesian markov chain monte carlo methods. *Ecology letters*, Wiley Online Library, v. 10, n. 7, p. 551–563, 2007.
- LELE, S. R.; NADEEM, K.; SCHMULAND, B. Estimability and likelihood inference for generalized linear mixed models using data cloning. *Journal of the American Statistical Association*, 2012.
- LUCAS, R. E. Econometric policy evaluation: A critique. In: ELSEVIER. *Carnegie-Rochester conference series on public policy*. [S.l.], 1976. v. 1, p. 19–46.
- MCFADDEN, D. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, p. 995–1026, 1989.
- MOTULA, P. F. N. Estimation of dsge models: A monte carlo analysis.
- NELDER, J. A.; MEAD, R. A simplex method for function minimization. *The computer journal*, Br Computer Soc, v. 7, n. 4, p. 308–313, 1965.
- PAKES, A.; POLLARD, D. Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, p. 1027–1057, 1989.
- ROTHENBERG, T. J. Identification in parametric models. *Econometrica: Journal of the Econometric Society*, p. 577–591, 1971.
- RUGE-MURCIA, F. J. Methods to estimate dynamic stochastic general equilibrium models. *Journal of Economic Dynamics and Control*, Elsevier, v. 31, n. 8, p. 2599–2636, 2007.

- SARGENT, T. J. The observational equivalence of natural and unnatural rate theories of macroeconomics. *The Journal of Political Economy*, p. 631–640, 1976.
- SIMS, C. Matlab optimization software. *QM&RBC Codes*, Quantitative Macroeconomics & Real Business Cycles, 1999.
- SIMS, C. A. Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, p. 1–48, 1980.
- SIMS, C. A. Bayesian skepticism on unit root econometrics. *Journal of Economic dynamics and Control*, Elsevier, v. 12, n. 2, p. 463–474, 1988.
- SIMS, C. A. Solving linear rational expectations models. *Computational economics*, Springer, v. 20, n. 1, p. 1–20, 2002.
- SMETS, F.; WOUTERS, R. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, Wiley Online Library, v. 1, n. 5, p. 1123–1175, 2003.
- SMETS, F.; WOUTERS, R. Shocks and frictions in us business cycles: A bayesian dsge approach. *The American Economic Review*, American Economic Association, v. 97, n. 3, p. 586–606, 2007.

APPENDIX A – Tables and Figures

Table 1 – Parameter Definition

Parameter	Description
ψ	Investment adjustment cost
σ_c	Inverse of the elasticity of intertemporal substitution
h	Consumption habit
ξ_w	Calvo wage reoptimization probability
σ_l	Elasticity of labor supply
ξ_p	Calvo price reoptimization probability
ι_w	Degree of wage indexation to past inflation
ι_p	Degree of price indexation to past inflation
Ψ	Capital utilization adjustment cost
ϕ	Fixed costs in production
r_π	Response of interest rates to past inflation
ρ	Monetary policy smoothing parameter
r_y	Response of interest rates to output
$r_{\Delta y}$	Response of interest rates to output gap
$\bar{\pi}$	Steady-state inflation rates
$100(\beta^{-1} - 1)$	Discount factors
$\bar{l}$	Steady-state worked hours
$100(\gamma - 1)$	Technology growth trend
$\rho_g a$	Effect of non-specific technology innovations on government spending
α	Steady-state share of capital in production
δ	Capital depreciation rate
g_y	Government spending to Output ratio
ϕ_w	Steady-state labor market mark-up
ϵ_w	Kimball (wage)
ϵ_p	Kimball (price)
ρ_a	Non-specific technology autorregressive parameter
ρ_b	Risk premium shock autorregressive parameter
ρ_g	Government spending autorregressive parameter
ρ_i	Investment-specific technology autorregressive parameter
ρ_m	Monetary policy shock autorregressive parameter
ρ_p	Price-inflation shock autorregressive parameter
ρ_w	Wage-inflation shock autorregressive parameter
μ_p	Price-inflation shock moving-average parameter
μ_w	Wage-inflation shock moving-average parameter
σ_a	Std Dev of non-specific technology innovations
σ_b	Std Dev of risk premium innovations
σ_g	Std Dev of government spending innovations
σ_i	Std Dev of investment-specific technology innovations
σ_m	Std Dev of monetary policy innovations
σ_p	Std Dev of price-inflation innovations
σ_w	Std Dev of wage-inflation innovations

Table 2 – Data cloning posterior estimates (1)

Parameter	Prior	Optimization			Single-Sample		5 Clones		10 Clones		25 Clones		Benchmark
	Density	Mean	Std	Mode	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean
ψ	N	4.00	1.50	5.49	5.588	2.241	6.851	1.091	6.004	0.694	5.974	1.044	5.48
σ_c	N	1.50	0.38	1.47	1.435	0.328	1.401	0.327	1.406	0.243	1.468	0.061	1.39
h	B	0.70	0.10	0.70	0.705	0.102	0.794	0.071	0.771	0.032	0.759	0.054	0.71
ξ_w	B	0.50	0.10	0.73	0.692	0.163	0.896	0.024	0.855	0.067	0.838	0.121	0.7
σ_l	N	2.00	0.75	1.67	1.655	1.259	3.791	0.433	3.267	0.731	3.398	1.852	1.83
ξ_p	B	0.50	0.10	0.68	0.676	0.125	0.784	0.048	0.737	0.039	0.695	0.076	0.66
ι_w	B	0.50	0.15	0.56	0.543	0.277	0.476	0.145	0.495	0.141	0.582	0.037	0.58
ι_p	B	0.50	0.15	0.24	0.26	0.202	0.273	0.092	0.276	0.073	0.265	0.039	0.23
Ψ	B	0.50	0.15	0.40	0.411	0.211	0.171	0.122	0.199	0.102	0.226	0.388	0.54
ϕ	N	1.25	0.12	1.65	1.643	0.171	1.581	0.095	1.591	0.065	1.575	0.078	1.6
r_π	N	1.50	0.25	1.98	2.015	0.375	2.131	0.316	1.936	0.072	1.949	0.193	2.04
ρ	N	0.75	0.10	0.82	0.82	0.052	0.872	0.016	0.854	0.018	0.851	0.021	0.81
r_y	N	0.13	0.05	0.09	0.095	0.052	0.138	0.069	0.123	0.029	0.118	0.007	0.08
$r_{\Delta y}$	N	0.13	0.05	0.22	0.222	0.064	0.191	0.013	0.185	0.027	0.192	0.047	0.22
$\bar{\pi}$	G	0.63	0.10	0.67	0.679	0.157	0.534	0.121	0.604	0.107	0.604	0.065	0.78
$100(\beta^{-1} - 1)$	G	0.25	0.10	0.21	0.241	0.203	0.232	0.167	0.186	0.041	0.195	0.038	0.16
$\bar{l}$	N	0.00	2.00	0.40	0.296	2.249	2.351	1.921	0.876	1.135	1.038	1.222	0.53
$100(\gamma - 1)$	N	0.40	0.10	0.44	0.435	0.035	0.451	0.011	0.456	0.025	0.455	0.011	0.43
ρ_{ga}	N	0.50	0.25	0.60	0.587	0.204	0.553	0.097	0.644	0.116	0.625	0.036	0.52
α	N	0.30	0.05	0.32	0.314	0.092	0.354	0.106	0.371	0.031	0.356	0.067	0.19

Table 3 – Data cloning posterior estimates (2)

Parameter		Prior		Optimization	Single-Sample		5 Clones		10 Clones		25 Clones		Benchmark
	Density	Mean	Std	Mode	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean
ρ_a	B	0.50	0.20	0.95	0.948	0.0354	0.974	0.027	0.968	0.014	0.967	0.005	0.95
ρ_b	B	0.50	0.20	0.17	0.212	0.1957	0.219	0.183	0.217	0.034	0.209	0.014	0.22
ρ_g	B	0.50	0.20	0.97	0.973	0.0196	0.997	0.001	0.996	0.016	0.995	0.019	0.97
ρ_i	B	0.50	0.20	0.74	0.747	0.134	0.647	0.083	0.671	0.101	0.681	0.045	0.71
ρ_m	B	0.50	0.20	0.12	0.142	0.1294	0.042	0.049	0.121	0.095	0.121	0.031	0.15
ρ_p	B	0.50	0.20	0.90	0.879	0.1326	0.999	0.001	0.999	0.075	0.999	0.016	0.89
ρ_w	B	0.50	0.20	0.97	0.959	0.0366	0.972	0.029	0.969	0.004	0.967	0.011	0.96
μ_p	B	0.50	0.20	0.77	0.723	0.2217	0.966	0.013	0.952	0.115	0.927	0.081	0.69
μ_w	B	0.50	0.20	0.87	0.809	0.1694	0.945	0.048	0.914	0.026	0.904	0.029	0.84
σ_a	IG	0.10	2.00	0.43	0.434	0.0506	0.472	0.029	0.445	0.011	0.446	0.026	0.45
σ_b	IG	0.10	2.00	0.24	0.237	0.0606	0.243	0.052	0.241	0.038	0.243	0.003	0.23
σ_g	IG	0.10	2.00	0.51	0.516	0.0519	0.52	0.016	0.504	0.061	0.512	0.011	0.53
σ_i	IG	0.10	2.00	0.43	0.435	0.0688	0.469	0.079	0.458	0.085	0.461	0.011	0.56
σ_m	IG	0.10	2.00	0.24	0.244	0.1014	0.231	0.011	0.233	0.006	0.228	0.009	0.24
σ_p	IG	0.10	2.00	0.14	0.141	0.0328	0.145	0.016	0.139	0.014	0.133	0.004	0.14
σ_w	IG	0.10	2.00	0.24	0.235	0.0359	0.231	0.024	0.217	0.008	0.224	0.028	0.24

Figure 1 – Single sample MCMC estimate posterior distributions. (1)

Gray solid line: prior distribution.

Black solid line: posterior distribution.

Dashed line: posterior mode.

Figure 2 – Single sample MCMC estimate posterior distributions. (2)

Figure 3 – Single sample MCMC estimate posterior distributions. (3)

Figure 4 – Single sample MCMC estimate posterior distributions. (4)

$\bar{\beta}$ denotes $100(\beta^{-1} - 1)$, and $\bar{\gamma}$ denotes $100(\gamma - 1)$.

Figure 5 – 25 Clones MCMC estimate posterior distributions. (1)

Gray solid line: prior distribution.

Black solid line: posterior distribution.

Dashed line: posterior mode.

Figure 6 – 25 Clones MCMC estimate posterior distributions. (2)

Figure 7 – 25 Clones MCMC estimate posterior distributions. (3)

Figure 8 – 25 Clones MCMC estimate posterior distributions. (4)

$\bar{\beta}$ denotes $100(\beta^{-1} - 1)$, and $\bar{\gamma}$ denotes $100(\gamma - 1)$.

Table 4 – Normalized data cloning posterior standard deviation

Parameters	s_1^*	s_2^*	s_3^*	s_5^*	s_{10}^*	s_{25}^*
Ψ	1.000	0.389	0.373	0.582	0.445	1.841
σ_l	1.000	0.441	0.348	0.344	0.33	1.471
ρ_g	1.000	0.178	0.142	0.081	0.132	1.004
σ_w	1.000	0.406	0.434	0.685	0.328	0.799
ξ_w	1.000	0.176	0.174	0.151	0.318	0.738
α	1.000	0.428	0.545	1.149	0.495	0.733
$r_{\Delta y}$	1.000	0.338	0.361	0.207	0.515	0.733
ξ_p	1.000	0.339	0.343	0.382	0.409	0.61
l	1.000	0.621	0.575	0.853	1.046	0.543
h	1.000	0.383	0.294	0.7	0.127	0.534
σ_a	1.000	0.284	0.533	0.588	0.792	0.519
r_π	1.000	0.352	0.571	0.843	0.225	0.516
ψ	1.000	0.429	0.423	0.486	0.206	0.466
ϕ	1.000	0.488	0.611	0.556	0.441	0.456
$\bar{\pi}$	1.000	0.529	0.375	0.763	1.059	0.413
ρ	1.000	0.299	0.329	0.316	0.193	0.397
μ_p	1.000	0.097	0.086	0.059	0.087	0.362
ρ_i	1.000	0.426	0.452	0.621	0.555	0.341
$100(\gamma - 1)$	1.000	0.314	0.257	0.311	0.226	0.311
ρ_w	1.000	0.265	0.295	0.803	0.532	0.306
ρ_m	1.000	0.326	0.333	0.383	0.912	0.238
σ_g	1.000	0.425	0.483	0.312	0.676	0.204
ι_p	1.000	0.485	0.288	0.455	0.433	0.194
σ_c	1.000	0.425	0.465	0.999	0.536	0.187
$100(\beta^{-1} - 1)$	1.000	0.705	0.525	0.825	0.226	0.186
ρ_{ga}	1.000	0.357	0.497	0.475	0.328	0.178
μ_w	1.000	0.108	0.109	0.285	0.324	0.171
σ_i	1.000	0.302	0.773	1.156	1.315	0.175
ρ_a	1.000	0.209	0.333	0.779	0.581	0.163
r_y	1.000	0.316	0.728	1.324	0.277	0.148
ι_w	1.000	0.503	0.384	0.523	0.628	0.135
σ_p	1.000	0.307	0.371	0.496	0.368	0.128
ρ_p	1.000	0.027	0.017	0.007	0.002	0.122
σ_m	1.000	0.359	0.116	0.116	0.146	0.088
ρ_b	1.000	0.392	0.386	0.937	0.325	0.072
σ_b	1.000	0.402	0.419	0.871	0.712	0.061

Figure 9 – Normalized standard error. (1)

Figure 10 – Normalized standard error. (2)

Figure 11 – Normalized standard error. (3)

Figure 12 – Normalized standard error. (4)

Figure 13 – Standardized maximum eigenvalue of the posterior covariance matrix

APPENDIX B – The log-linearized Smets and Wouters (2007) model

B.1 The Sticky Price and Wage Equilibrium Conditions

The log-linearized aggregate resource constraint of this economy is given by

$$\hat{y}_t = c_y \hat{c}_t + i_y \hat{i}_t + z_y \hat{z}_t + \varepsilon_t^g, \quad (\text{B.1})$$

where $\hat{y}_t$ is real GDP. It is composed by real private consumption ($\hat{c}_t$), real private investment ($\hat{i}_t$), the capital utilization rate ($\hat{z}_t$), and exogenous spending (ε_t^g). The parameter c_y is the steady-state consumption-output ratio, i_y is the steady-state investment-output ratio, and z_y is the steady-state capital income, where

$$c_y = 1 - i_y - g_y, \quad (\text{B.2})$$

$$i_y = (\gamma + \delta - 1)k_y, \quad (\text{B.3})$$

$$z_y = r^k k_y. \quad (\text{B.4})$$

The steady-state exogenous spending-output ratio is given by g_y , δ is the capital depreciation rate, the steady-state capital-output ratio is represented by k_y , and γ is the steady-state growth rate.

Consumption dynamics follows from the consumption Euler equation,

$$\hat{c}_t = c_1 \hat{c}_{t-1} + (1 - c_1) E_t \hat{c}_{t+1} + c_2 (\hat{l}_t - E_t \hat{l}_{t+1}) - c_3 (\hat{r}_t - E_t \hat{\pi}_{t+1}) + \varepsilon_t^b, \quad (\text{B.5})$$

where $\hat{l}_t$ is worked hours, $\hat{r}_t$ is the nominal interest rate controlled by the monetary authority, and ε_t^b is an exogenous preferences shifting process. The three parameters of the consumption Euler equation are:

$$c_1 = \frac{\lambda\gamma}{1 + (\lambda\gamma)}, \quad c_2 = \frac{(\sigma_c - 1)(w^h l/c)}{\sigma_c(1 + (\lambda\gamma))}, \quad c_3 = \frac{1 - (\lambda\gamma)}{\sigma_c(1 + (\lambda\gamma))}. \quad (\text{B.6})$$

External habit formation is measured by the parameter λ , i.e. measures how much past consumption matters to current consumption decision. σ_c is the inverse of the elasticity of inter-temporal substitution for constant labor, and w^hl/c is the steady-state hourly real wage bill to consumption ratio. Reckon that if $\sigma_c = 1$ (log-utility) and $\lambda=0$ (no external habit), then (B.5) is reduced to a purely forward looking form.

Investment evolution over time is similarly described by

$$\hat{i}_t = i_1 \hat{i}_{t-1} + (1 - i_1) E_t \hat{i}_{t+1} + i_2 \hat{q}_t + \varepsilon_t^i, \quad (\text{B.7})$$

where $\hat{q}_t$ is the real value of installed capital, and ε_t^i is an investment-specific technology exogenous process. The parameters of (B.7) are given by

$$i_1 = \frac{1}{1 + \beta\gamma(1 - \sigma_c)}, \quad i_2 = \frac{1}{(1 + \beta\gamma(1 - \sigma_c))\gamma^2\varphi},$$

where β is the discount factor used by households, and φ is the steady-state elasticity of the capital adjustment cost function.

The value of the capital stock, $\hat{q}_t$, is determined so that ¹

$$\hat{q}_t = q_1 E_t \hat{q}_{t+1} + (1 - q_1) E_t \hat{r}_{t+1}^k - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + c_3^{-1} \varepsilon_t^b, \quad (\text{B.8})$$

where $\hat{r}_t^k$ is the rental rate of capital. The parameter of (B.8) is

$$q_1 = \beta\gamma^{-\sigma_c}(1 - \delta) = \frac{1 - \delta}{r^k + 1 - \delta}$$

Now concerning the supply-side of the economy, the log-linearized aggregate production function is given by

$$\hat{y}_t = \phi_p [\alpha \hat{k}_t^s + (1 - \alpha) \hat{l}_t + \varepsilon_t^a], \quad (\text{B.9})$$

where $\hat{k}_t^s$ regards capital services used in production, and ε_t^a is an exogenous total factor productivity process. The parameter α reflects the share of capital in production, while ϕ_p is equal to one plus the steady-state share of fixed costs in production.

¹ It seems important to clarify the equity premium shock in SW07 is actually a transformation of the exogenous preference shifting process, it is $\varepsilon^{equitypremium} = c_3^{-1} \varepsilon^b$. This becomes evident when inspecting the author's Dynare code. Thus both because this presentation follows their code and for simplicity, we chose not to treat equity premium innovations as a distinct process.

The capital services variable introduces a "time to build" dynamic to the model, i.e. newly installed capital only becomes effective with a one period lag. This means that

$$\hat{k}_t^s = \hat{k}_{t-1} + \hat{z}_t, \quad (\text{B.10})$$

where $\hat{k}_t$ is the installed capital. The degree of capital utilization, $\hat{z}_t$, is a positive function of the rental rate of capital so that

$$\hat{z}_t = z_1 \hat{r}_t^k, \quad (\text{B.11})$$

where

$$z_1 = \frac{1 - \psi}{\psi},$$

and ψ is a positive function of the elasticity of the capital adjustment function and normalized to be between 0 and 1. The larger ψ is, the costlier it is to change capital utilization.

The log-linearized installed capital motion equation is

$$\hat{k}_t = k_1 \hat{k}_{t-1} + (1 - k_1) \hat{i}_t + k_2 \varepsilon_t^i, \quad (\text{B.12})$$

and the two parameters are given by

$$k_1 = \frac{1 - \delta}{\gamma}, \quad k_2 = (\gamma + \delta - 1)(1 + \beta\gamma^{-\sigma_c})\gamma\psi$$

From the monopolistically competitive goods market, price markup ($\hat{\mu}_t^p$) is equal to minus the real marginal cost ($\hat{\mu}_t^c$) under cost maximization by firms. That is,

$$\hat{\mu}_t^p = \alpha(\hat{k}_t^s - \hat{l}_t) - \hat{w}_t + \varepsilon_t^a, \quad (\text{B.13})$$

where the real wage is $\hat{w}_t$. Similarly, the real marginal cost is

$$\hat{\mu}_t^c = \alpha \hat{r}_t^k + (1 - \alpha) \hat{w}_t - \varepsilon_t^a, \quad (\text{B.14})$$

notice that (B.14) is obtained by substituting for the optimally determined capital-labor ratio in (B.16). where σ_l is the elasticity of labor supply with respect to the real wage.

Profit maximization by price-setting firms yields the log-linearized price Phillips curve

$$\begin{aligned}\hat{\pi}_t &= \pi_1 \hat{\pi}_{t-1} + \pi_2 E_t \hat{\pi}_{t-1} - \pi_3 \hat{\mu}_t^p + \varepsilon_t^p \\ &= \pi_1 \hat{\pi}_{t-1} + \pi_2 E_t \hat{\pi}_{t-1} + \pi_3 \hat{\mu}_t^c + \varepsilon_t^p,\end{aligned}\tag{B.15}$$

where ε_t^p is an exogenous price markup process. The parameters of (B.15) are so that

$$\pi_1 = \frac{\iota_p}{1 + \beta\gamma^{1-\sigma_c}\iota_p}, \pi_2 = \frac{\beta\gamma^{1-\sigma_c}}{1 + \beta\gamma^{1-\sigma_c}\iota_p}, \pi_3 = \frac{(1 - \xi_p)(1 - \beta\gamma^{1-\sigma_c}\xi_p)}{(1 + \beta\gamma^{1-\sigma_c}\iota_p)\xi_p((\phi_p - 1)\epsilon_p + 1)}.$$

The degree of indexation to past inflation is determined by the parameter ι_p , ξ_p measures the degree of price stickiness such that $1 - \xi_p$ is the probability that a firm can be reoptimize its price, and ε_p is the curvature of the Kimball (1995) goods market aggregator.

$$\hat{r}_t^k = -(\hat{k}_t^s - \hat{l}_t) + \hat{w}_t\tag{B.16}$$

The wage markup households set in the monopolistically competitive labor market is equal to the difference between the real wage and the marginal rate of substitution between labor and consumption

$$\hat{\mu}_t^w = \hat{w}_t - \left[\sigma_l \hat{l}_t + \frac{1}{1 - (\lambda\gamma)} \left(\hat{c}_t - \frac{\lambda}{\gamma} \hat{c}_{t-1} \right) \right]\tag{B.17}$$

As real wages are sticky, following a Calvo (1983) structure, real wage respond only gradually to the desired wage markup

$$\hat{w}_t = w_1 \hat{w}_{t-1} + (1 - w_1)[E_t \hat{w}_{t+1} + E_t \hat{\pi}_{t+1}] - w_2 \hat{\pi}_t + w_3 \hat{\pi}_{t-1} - w_4 \hat{\mu}_t^w + \varepsilon_t^w,\tag{B.18}$$

where ε_t^w is an exogenous wage markup process. The parameters of (B.18) are

$$\begin{aligned}w_1 &= \frac{1}{1 + \beta\gamma^{1-\sigma_c}}, & w_2 &= \frac{1 + \beta\gamma^{1-\sigma_c}\iota_w}{1 + \beta\gamma^{1-\sigma_c}}, \\ w_3 &= \frac{\iota_w}{1 + \beta\gamma^{1-\sigma_c}}, & w_4 &= \frac{(1 - \xi_w)(1 - \beta\gamma^{1-\sigma_c}\xi_w)}{(1 + \beta\gamma^{1-\sigma_c})\xi_w((\phi_w - 1)\epsilon_w + 1)}.\end{aligned}$$

Wage indexation degree to past inflation is given by ι_w , while ξ_w is the degree of wage stickiness. The steady-state labor markup is given by $\phi_w - 1$ and ϵ_w is the curvature of the Kimball (1995) labor market aggregator.

The monetary authority behaves accordingly to the monetary policy reaction function

$$\hat{r}_t = \rho \hat{r}_{t-1} + (1 - \rho)[r_\pi \hat{\pi}_t + r_y(\hat{y}_t - \hat{y}_t^f)] + r_{\Delta y}[\Delta \hat{y}_t - \Delta \hat{y}_t^f] + \varepsilon_t^r, \quad (\text{B.19})$$

where $\hat{y}_t^f$ is the potential output, measured as the flexible prices output.

B.2 The Flexible Price and Wage Equilibrium Conditions

The flexible price equations can be obtained by setting two exogenous wage markup processes, μ_t^p and μ_w , equal to zero; and eliminating wage and price stickiness, i.e. $\xi_w = \xi_p = 0$, and $\iota_w = \iota_p = 0$. Equations (B.20a) to (B.20k) describe the flexible prices version of this economy:

$$\hat{y}_t^f = c_y + \hat{c}_t^f + i_y \hat{y}_t^f + z_y \hat{z}_t^f + \varepsilon_t^g, \quad (\text{B.20a})$$

$$\hat{c}_t^f = c_1 \hat{c}_{t-1}^f + (1 - c_1) E_t \hat{c}_{t+1}^f + c_2 (\hat{l}_t^f - E_t \hat{l}_{t+1}^f) - c_3 \hat{r}_t^f + \varepsilon_t^b, \quad (\text{B.20b})$$

$$\hat{i}_t^f = i_1 \hat{i}_{t-1}^f + (1 - i_1) E_t \hat{i}_{t+1}^f + i_2 \hat{q}_t^f + \varepsilon_t^i, \quad (\text{B.20c})$$

$$\hat{q}_t^f = q_1 E_t \hat{q}_{t+1}^f + (1 - q_1) E_t \hat{r}_{t+1}^{k,f} - \hat{r}_t^f + c_3^{-1} \varepsilon_t^b, \quad (\text{B.20d})$$

$$\hat{y}_t^f = \phi_p [\alpha \hat{k}_t^{s,f} + (1 - \alpha) \hat{l}_t^f + \varepsilon_t^a], \quad (\text{B.20e})$$

$$\hat{k}_t^{s,f} = \hat{k}_{t-1}^f + \hat{z}_t^f, \quad (\text{B.20f})$$

$$\hat{z}_t^f = z_1 \hat{r}_t^{k,f} \quad (\text{B.20g})$$

$$\hat{k}_t^f = k_1 \hat{k}_{t-1}^f + (1 - k_1) \hat{i}_t^f + k_2 \varepsilon_t^i, \quad (\text{B.20h})$$

$$\varepsilon_t^a = \alpha \hat{r}_t^{k,f} + (1 - \alpha) \hat{w}_t^f, \quad (\text{B.20i})$$

$$\hat{r}_t^{k,f} = -(\hat{r}_t^{k,f} - \hat{l}_t^f) + \hat{w}_t^f, \quad (\text{B.20j})$$

$$\hat{w}_t^f = \sigma_l \hat{l}_t^f + \frac{1}{1 - (\lambda \gamma)} [\hat{c}_t^f - \frac{\lambda}{\gamma} \hat{c}_{t-1}^f], \quad (\text{B.20k})$$

in which $\hat{r}_t^f$ is the real interest rate that would prevail were prices and wages flexible.

B.3 The Exogenous Variables

The seven innovation processes are considered in sw07 are described. Five are specified as AR(1) processes: government spending², ε_t^g ; preferences shifting ε_t^b , investment-specific technology, ε_t^i ; non-specific technology ε_t^a , and monetary policy ε_t^r . The remaining two innovation, wage markup ε_t^w and price inflation ε_t^p , are characterized by ARMA(1,1) processes. It is

$$\varepsilon_t^g = \rho_g \varepsilon_{t-1}^g + \sigma_g \eta_t^g + \rho_{ga} \sigma_a \eta_a, \quad (\text{B.21a})$$

$$\varepsilon_t^b = \rho_b \varepsilon_{t-1}^b + \sigma_b \eta_t^b, \quad (\text{B.21b})$$

$$\varepsilon_t^i = \rho_i \varepsilon_{t-1}^i + \sigma_i \eta_t^i, \quad (\text{B.21c})$$

$$\varepsilon_t^a = \rho_a \varepsilon_{t-1}^a + \sigma_a \eta_t^a, \quad (\text{B.21d})$$

$$\varepsilon_t^p = \rho_p \varepsilon_{t-1}^p + \sigma_p \eta_t^p - \mu_p \sigma_p \eta_{t-1}^p, \quad (\text{B.21e})$$

$$\varepsilon_t^w = \rho_w \varepsilon_{t-1}^w + \sigma_w \eta_t^w - \mu_w \sigma_w \eta_{t-1}^w, \quad (\text{B.21f})$$

$$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + \sigma_r \eta_t^r, \quad (\text{B.21g})$$

where the shocks $\eta_t^j, j = a, b, g, i, p, r, w$ are i.i.d. $N(0, 1)$ random variables.

B.4 The Steady-State Relations

Steady-state quantities have already been provided for the consumption-output ratio (B.2), investment-output ratio (B.3), and capital income (B.4). Thus it remains to provide expressions for the steady-state values of the rental rate of capital B.22, real wage (B.23), labor-capital ratio (B.24), capital-output ratio (B.25), and hourly real wage bill

² Note that the exogenous spending process depends both on the exogenous spending shock, η_t^g , and the contemporary non-specific technology shock, η_t^a .

to consumption ratio (B.27).

$$r^k = \frac{1}{\beta\gamma^{-\sigma_c} + \delta - 1}, \quad (\text{B.22})$$

$$w = \left[\frac{\alpha^\alpha (1 - \alpha)^{(1-\alpha)}}{\phi_p (r^k)^\alpha} \right]^{1(1-\alpha)}, \quad (\text{B.23})$$

$$\frac{l}{k} = \frac{(1 - \alpha)r^k}{\alpha w}, \quad (\text{B.24})$$

$$\frac{k}{y} = \phi_p \left[\frac{l}{k} \right]^{\alpha-1}. \quad (\text{B.25})$$

$$(\text{B.26})$$

From (B.23)-(B.25) it is possible to show that $z_y = r^k k_y = \alpha, k_y = ky$.

The steady-state relation between real wages and hourly real wage and hourly real wage is $w = \phi_w w^h$. Thus the steady-state hourly wage bill to consumption ratio is given by

$$\frac{w^h l}{c} = \frac{(1 - \alpha)r^k k_y}{\phi_w \alpha c_y} = \frac{1 - \alpha}{\phi_w c_y}. \quad (\text{B.27})$$