GRIPS Discussion Paper 24-6

Estimation of Nonlinear DSGE Models Through Laplace Based Solutions

By

Elnura Baiaman kyzy
Roberto Leon-Gonzalez

July 2024

National Graduate Institute for Policy Studies
7-22-1 Roppongi, Minato-ku,
Tokyo, Japan 106-8677

Estimation of Nonlinear DSGE Models Through Laplace Based Solutions*

Elnura Baiaman kyzy
HIAS, Hitotsubashi University

Roberto Leon-Gonzalez
National Graduate Institute for Policy Studies (GRIPS)

July, 2024

*We gratefully acknowledge helpful comments from Gianni Amisano, Junichi Fujimoto, Hidehiko Matsumoto, as well as from participants at the 2022 NBER-NSF Seminar on Bayesian Inference in Econometrics and Statistics (SBIES), the 12^{th} European Seminar on Bayesian Econometric (ESOBE), 14^{th} RCEA Bayesian Econometrics Workshop and EcoSta 2024 conference. We also acknowledge financial support from the Japan Society for the Promotion of Science (19K01588) and from GRIPS Policy Research Center (P223RP214). All errors are our own.

Abstract

This paper proposes a novel Laplace based solution to nonlinear DSGE models that has a closed form likelihood. We implicitly use a nonlinear approximation to the policy function that is invertible with respect to the shocks, implying that in the approximation the shocks can be recovered uniquely from some of the control variables. Using perturbation methods and a Lagrange inversion formula we are able to calculate the derivatives of the likelihood and construct the Laplace based solution. In contrast with previous likelihood-based approaches, the method used here requires neither the introduction of linear shocks nor simulation to evaluate the likelihood. Using US data we estimate linear and nonlinear variants of a well-known neoclassical growth model with and without time-varying variances. We find that a nonlinear heteroscedastic model has a much better empirical performance. Furthermore, our models allow us to ascertain that the monetary policy shock causes 95% of the time changes in economic uncertainty.

JEL Classification: E0, C63

Keywords: Economic Uncertainty, Time-Varying Volatility, Risk-Premium, Higher-Order Approximation.

1 Introduction

Linearization of Dynamic Stochastic General Equilibrium (DSGE) models is a common tool to approximate the solution of the dynamic optimization problem in a DSGE (e.g., Blanchard and Kahn (1985), Sims (2002), Klein (2000)). Linearization is typically achieved by using only the first term of a Taylor expansion around the steady state of the log of the equations representing the first-order conditions of the dynamic optimization problem.

Linearization made possible the use of formal statistical methods to estimate and test DSGE models, such as the General Method of Moments (e.g., Hansen and Singleton (1983), Christiano and Eichenbaum (1992)), Maximum Likelihood (e.g., Hansen and Sargent (1980), Altug (1989)) and Bayesian methods (e.g., DeJong et al. (2000), Schorfheide (2000), Otrok (2001)).

However, the linearization of this class of models also comes at a cost, because not all questions can be fully addressed in a linearized model. As argued by Schmitt-Grohé and Uribe (2004), two such questions are welfare evaluations and risk premia in stochastic environments. In a linearized model, the agents become risk-neutral, so it is impossible to analyze the impact of uncertainty on the economy. Furthermore, evaluating social welfare across alternative stochastic policy environments using a linear approximation model leads to the omission of some critical second-order terms, resulting in spurious results.

To address the limitations of the linearization method, Schmitt-Grohé and Uribe (2004) proposed the use of perturbation methods (e.g., Judd (1998)) to obtain higher-order Taylor approximations of the policy functions. They also showed that when perturbation methods are used to obtain a first-order approximation, the result is the same as that obtained in the previous literature on linearization (e.g., Blanchard and Kahn (1985)).

In terms of estimation, moving away from a linearized model towards a nonlinear one is challenging because the structural errors enter the model in a nonlinear fashion, and the likelihood function is no longer normal or readily available. Despite this difficulty, the seminal paper of Amisano and Tristani (2011) shows how to obtain the likelihood in some restricted models when using a second-order approximation to solve the model. The method is applicable only when there are no unobserved non-stochastic state variables, which implies for example that in a model with capital, capital has to be an observed variable. In addition, this method requires finding solutions of polynomial equations, which is computationally intensive, and therefore in practice the number of structural shocks has to be relatively small.

In a groundbreaking paper, Fernández-Villaverde and Rubío-Ramirez (2007) propose a particle filter approach that permits the numerical approximation of the likelihood and Bayesian estimation. However, because evaluating the likelihood requires simulation with a potentially large number of particles, the estimation method is slow, and impracticable in moderately large models. Furthermore, this approach requires the number of shocks to be greater than the number of observed variables and that a number of shocks (structural or measurement errors) enter linearly in the likelihood.

Amisano and Tristani (2011) show that when a DSGE model is approximated with higher order perturbation methods, the likelihood of the approximated model can be derived in theory, but is often numerically intractable. To surmount this problem, we solve the DSGE model in an alternative manner, proposing a higher order Laplace based solution to the DSGE, whose likelihood can be easily evaluated in closed form. We build on the perturbation methods literature to find out the mode and higher order derivatives of the likelihood that form the basis of the Laplace based solution of the DSGE model. The likelihood of this DSGE model solution is tractable and can be used in Bayesian or Maximum Likelihood estimation. In contrast with Amisano and Tristani (2011), this method allows for unobserved non-stochastic state variables, and unlike Fernández-Villaverde and Rubío-Ramirez (2007) it requires neither the introduction of linear shocks nor simulation to evaluate the likelihood.

A limitation of our method is that the likelihood can only be evaluated in closed form if the number of stochastic shocks and the number of observed variables are equal. However, since the researcher has often some flexibility in choosing these, our method can be used to estimate a wide range of models.

Using US data, we use our novel approach to estimate the well-known neoclassical growth model of Fernández-Villaverde (2010). In addition to the baseline model, we extend the model by allowing the shocks to be heteroscedastic in two different manners: i) using independent GARCH processes and ii) using a novel common factor GARCH process that implies that all volatilities move in unison. We find that the nonlinear DSGE extended with a common factor GARCH is much superior to the linear DSGE and all other nonlinear variants in terms of the predictive and marginal likelihoods. Furthermore, the common factor GARCH specification allows us to determine that 95%

of the time-variation in volatilities are caused by the monetary policy shock.

The remainder of this paper is structured as follows. Section 2 describes the perturbation method that has been used in previous literature to obtain the policy function and for estimation. Section 3 presents the Laplace based solution of the DSGE model and its likelihood. Section 4 illustrates that the Laplace based solution to the model that we propose gives similar results to the previous approach in the literature. Then, section 5 uses the proposed method to estimate a neoclassical growth DSGE model and discusses the results. Finally, section 6 concludes.

2 Using Perturbation Methods to Obtain the Policy Function and Estimate the Model

Using the first-order conditions of the dynamic optimization problem, we can see that the general nonlinear form of a DSGE model can be cast as (e.g., Amisano and Tristani (2011)):

$$E_t[f(y_{t+1}, y_t, x_{t+1}, x_t)] = 0 \tag{2.1}$$

where E_t is the expectation operator conditional on information available at time t; y_t represents a vector of non-predetermined variables; and x_t denotes a vector of predetermined variables. The vector x_t can be partitioned as $x_t = (K_t', A_t')'$, where K_t is a vector of endogenous predetermined state variables and A_t is a vector of exogenous predetermined state variables. Schmitt-Grohé and Uribe (2004) assume that A_{t+1} follows the stochastic process: $A_{t+1} = \Lambda A_t + \sigma\varepsilon_t$, where the scalar σ is a perturbation parameter, and ε_t is a vector of zero mean innovations, independently and identically distributed with variance-covariance matrix Σ. The matrix Λ has all eigenvalues within the unit circle. The solution of the DSGE model consists of the policy functions which give the optimal value of y_t given x_t:

$$\begin{aligned} y_t &= g_y(K_t, A_t, \sigma) \\ K_{t+1} &= h_K(K_t, A_t, \sigma) \\ A_{t+1} &= h_A(K_t, A_t, \sigma) + \sigma\varepsilon_t \end{aligned} \tag{2.2}$$

Schmitt-Grohé and Uribe (2004) obtain a Taylor approximation of the above policy functions, g_y, h_K and h_A around the deterministic steady state, $x_t = x_{ss}$ and $\sigma = 0$, using perturbation methods. The deterministic steady state is defined as vectors (y_{ss}, x_{ss})

such that $f(y_{ss}, y_{ss}, x_{ss}, x_{ss}) = 0$. Perturbation methods (e.g., Fleming (1971) and Judd (1998)) provide a higher order Taylor expansion with respect to the state variables x_t as well as the scale parameter σ. Schmitt-Grohé and Uribe (2004) presented a set of MATLAB programs designed to compute the coefficients of the higher-order approximations. A similar approach was proposed by Sims (2000) and Collard and Juillard (2001).

In order to write the higher-order approximation around the deterministic steady state, let us define the vectors Y_t and X_t as:

$$Y_t = \begin{pmatrix} y_t - y_{ss}, \\ K_{t+1} - K_{ss} \\ A_{t+1} - A_{ss} \end{pmatrix}, \qquad X_t = \begin{pmatrix} K_t - K_{ss} \\ A_t - A_{ss} \end{pmatrix},$$

where y_{ss}, K_{ss}, A_{ss} are the deterministic steady state values of y_t, K_t, A_t. Using this notation the solution in (2.2) can be written as:

$$Y_t = g(X_t, \varepsilon_t, \sigma) \tag{2.3}$$

Following Dynare (2021), higher order approximations to (2.3) can be conveniently written using Kronecker products, and for example a second order approximation becomes as follows:

$$Y_t = G^{0,0} + G^{1,0}\varepsilon_t + G^{0,1}X_t + G^{2,0}(\varepsilon_t \otimes \varepsilon_t) + G^{0,2}(X_t \otimes X_t) + G^{1,1}(\varepsilon_t \otimes X_t) \tag{2.4}$$

where $G^{0,0}, \ldots, G^{1,1}$ are matrices of coefficients that depend on σ and other parameters of the model.

Let Y_t^o be the variables of Y_t for which there are observed data, for $t = 1, ..., T$. Taking as given the model solution in (2.4), Amisano and Tristani (2011) derive the likelihood when Y_t^o and ε_t have the same dimension n_o, but show that the method is only computationally feasible when K_t is observed or absent from the model. However, even in that case, this method requires finding all the 2^{n_o} solutions of polynomial equations for each $t = 1, ..., T$, so in practice it can only be used when n_o is small. For example, Amisano and Tristani (2011) provide an empirical application in which the dimension of the structural errors ε_t is two.

In a seminal paper, Fernández-Villaverde and Rubío-Ramirez (2007) propose a particle-filtering approach that permits the numerical approximation of the likelihood in nonlinear DSGE models. This method has the advantage that can be used in conjunction with several solution methods, including not only perturbation methods but also global solution methods, etc. (e.g., value function iteration). However, the method is slow because it uses simulation to evaluate the likelihood, and can become impracticable in models of moderate sizes. Furthermore, this method does not apply to (2.4) unless it is assumed that either: 1) Y_t^o is observed with measurement error; or 2) some elements of ε_t enter linearly in (2.4). Either of these linearity assumptions are necessary to calculate the importance weights in the particle filter. Although adding measurement errors might be justified and in some cases improves the empirical fit of the model, we think it is of interest to estimate the theoretical model with only the structural errors, which in most cases enter nonlinearly.

Kollmann (2017) proposed a novel method to obtain a tractable likelihood in the estimation of nonlinear DSGE models. However, in this method, nonlinear terms such as $(\varepsilon_t \otimes \varepsilon_t)$ are replaced by their unconditional expected value.

3 A Laplace Based Solution and the Likelihood

Writing the vector Y_t as $Y_t = ((Y_t^o)', (Y_t^n)')'$, the solution in (2.3) can be written as:

$$\begin{aligned} Y_t^o &= g_o(X_t, \varepsilon_t, \sigma) \\ Y_t^n &= g_n(X_t, \varepsilon_t, \sigma) \end{aligned} \tag{3.1}$$

We assume that Y_t^o and ε_t have the same dimension n_o and that both are defined in $\mathbb{R}^{n_o}$. If the function $g_o(X_t, \varepsilon_t, \sigma)$ was globally invertible with respect to ε_t in an area of probability 1, then the following inverses would be well-defined:

$$\begin{aligned} \varepsilon_t &= m_o(X_t, Y_t^o, \sigma) \\ Y_t^n &= m_n(X_t, Y_t^o, \sigma) \end{aligned} \tag{3.2}$$

Following Galewski (2016), the following conditions are sufficient for global invertibility:

1. $g_o(X_t, \varepsilon_t, \sigma)$ is differentiable in ε_t with continuous derivatives.

2. The determinant of the Jacobian of $g_o(X_t, \varepsilon_t, \sigma)$ with respect to ε_t is never 0.

3. $\|g_o(X_t, \varepsilon_t, \sigma)\| \to \infty$ as $\|\varepsilon_t\| \to \infty$, where $\|.\|$ is the norm operator.

The second condition is guaranteed if the policy function is strictly monotonic in ε_t. There is a large literature which gives conditions for policy functions to be strictly monotonic and proves the condition holds in many important models, for example: Topkis (1978), Hopenhayn and Prescott (1992), Stokey et al. (1989), Gordon and Qiu (2018). In addition, the assumption of monotonicity has often been exploited in the literature on DSGE models to obtain the policy function more efficiently (e.g. Christiano (1990), Judd (1998), Gordon and Qiu (2018)).

We assume that the policy function $Y_t^o = g_o(X_t, \varepsilon_t, \sigma)$ is *locally*, but not necessarily *globally*, invertible with respect to ε_t at the steady state. This condition holds if the Jacobian of the transformation from Y_t^o to ε_t is not zero at the steady state, which is a condition that can be verified numerically. This will allow us to obtain the Lagrange inverse of the Taylor approximation of $g_o(X_t, \varepsilon_t, \sigma)$[1]. In addition, we assume that there exists a globally invertible approximation of the policy function $Y_t^o = g_o(X_t, \varepsilon_t, \sigma)$ and we denote it with $Y_t^o = \hat{g}_o(X_t, \varepsilon_t, \sigma)$, and its inverse as $\varepsilon_t = \hat{m}_o(X_t, Y_t^o, \sigma)$. If the policy function was indeed globally invertible, then we would use $\hat{g}_o(X_t, \varepsilon_t, \sigma) = g_o(X_t, \varepsilon_t, \sigma)$. Otherwise $\hat{g}_o(X_t, \varepsilon_t, \sigma)$ is an invertible function that has the same Taylor polynomial as $g_o(X_t, \varepsilon_t, \sigma)$ up to a high order at the steady state. Two functions will have the same Taylor polynomial up to some order r, if they have the same value for the derivatives of order up to r at the point of approximation. Although there is a literature that constructs invertible approximations of functions (for example approximations of the cumulative density function of a normal distribution (e.g. Lipoth et al. (2022)), or approximations of multivariate functions with invertible neural networks (Teshima et al. (2020), Ishikawa et al. (2022)), for our purposes we do not need to obtain the invertible approximation explicitly: it is enough to assume that it exists.

We can use $\hat{m}_o(X_t, Y_t^o, \sigma)$ to obtain an approximation of the density of Y_t^o conditional on X_t, if we apply a change of variables theorem (e.g. Billingsley (1999))

$$\pi_y(Y_t^o|X_t) = \pi_\varepsilon(\hat{m}_o(X_t, Y_t^o, \sigma)) \left| \frac{\partial \hat{m}_o(X_t, Y_t^o, \sigma)}{\partial Y_t^o} \right| \tag{3.3}$$

where $\partial \hat{m}_o(X_t, Y_t^o, \sigma)/\partial Y_t^o$ is the Jacobian of the transformation and π_ε is the density

[1] Although Taylor polynomials are not in general globally invertible, the Lagrange inverse can be obtained and it approximates the inverse in a neighbourhood of the point of approximation.

function of ε.

The density in (3.3) can be approximated using a Laplace approximation of order M, which we denote as $\hat{\pi}_{(y,M)}(Y_t^o|X_t)$. A higher order Laplace approximation is a tractable density function which in logs has derivatives up to order M at the mode that coincide with those of the log objective density. Using the chain rule, in order to obtain the Laplace approximation, we only need the derivatives of $\varepsilon_t = \hat{m}_o(X_t, Y_t^o, \sigma)$ at the mode. Therefore we do not need to obtain the function $\varepsilon_t = \hat{m}_o(X_t, Y_t^o, \sigma)$ explicitly, only its derivatives. We approximate the mode of (3.3) and obtain the derivatives at the mode using the Taylor polynomial approximation of $\hat{m}_o(X_t, Y_t^o, \sigma)$ obtained through standard perturbation methods (as provided by Dynare (2021)) plus a Lagrange inversion formula to invert the Taylor polynomial.

The Laplace approximation of order M can be constructed using the following Polynomial-Normal (Skoulakis (2019)) density function:

$$\frac{1}{c}\exp\left(-\frac{1}{2}(Y-\mu)'\Omega^{-1}(Y-\mu)\right)\left(1+\sum_{j=3}^{M} B_j (Y-\mu)^{[j]}\frac{1}{j!}\right)^2 \tag{3.4}$$

where $x^{[j]}$ is the Kronecker power defined as:

$$x^{[2]} = x \otimes x, \qquad x^{[j]} = x^{[j-1]} \otimes x$$

and c is a normalizing constant. If $M = 2$ (such that $B_j = 0$ for all j) then (3.4) is just a normal density with mean μ and variance-covariance matrix Ω. In this case we obtain the standard Laplace approximation, with μ equal to the mode of the objective density and Ω equal to minus the inverse of the Hessian of the log objective density at the mode. For $M > 2$ this continues to be true, but the matrix $2B_j$ contains the j^{th} order derivatives of the log objective density at the mode. The normalizing constant c is known because all the moments of the normal density can be calculated in closed form.

Provided that $Y_t^o = g_o(X_t, \varepsilon_t, \sigma)$ is locally invertible at the steady state, its Taylor polynomial can be locally inverted using a Lagrange inversion formula. As a matter of notation, let the Lagrange inverse of the s order Taylor polynomial of $(Y_t^o = g_o(X_t, \varepsilon_t, \sigma), Y_t^n = g_n(X_t, \varepsilon_t, \sigma))$ be denoted as $(Y_t^o = \tilde{m}_{o,s}(X_t, Y_t^o, \sigma), Y_t^n = \tilde{m}_{n,s}(X_t, Y_t^o, \sigma))$. Specific formulas for these inversions are provided in Proposition 3.3 for the case $s = 2$. Then the derivatives of the likelihood in (3.3) can be approximated by those of the following function:

$$\tilde{\pi}_{(y,s)}(Y_t^o|X_t) = \pi_\varepsilon(\tilde{m}_{o,s}(X_t, Y_t^o, \sigma)) \left| \frac{\partial \tilde{m}_{o,s}(X_t, Y_t^o, \sigma)}{\partial Y_t^o} \right| \tag{3.5}$$

Because the Taylor polynomial $\tilde{m}_{o,s}(X_t, Y_t^o, \sigma)$ need not be globally invertible, the function $\tilde{\pi}_{(y,s)}(Y_t^o|X_t)$ in (3.5) is not guaranteed to be a proper density function, in the sense that the area under the curve does not need to add up to one. For this reason it cannot be used as an approximation for the likelihood. However, because the derivatives of the Taylor polynomial $\tilde{m}_{o,s}(X_t, Y_t^o, \sigma)$ approximate those of $\hat{m}_o(X_t, Y_t^o, \sigma)$ near the point of approximation, we can use the derivatives of $\tilde{\pi}_{(y,s)}(Y_t^o|X_t)$ in (3.5) to approximate the derivatives of the likelihood $\pi_y(Y_t^o|X_t)$ in (3.3). We can then use these derivatives to construct the Laplace approximation, which is a proper density in the sense that it integrates up to one.

We therefore propose to obtain the approximation $\hat{\pi}_{(y,s)}(Y_t^o|X_t)$ to the likelihood $\pi_y(Y_t^o|X_t)$ using the following procedure.

1. Obtain the Taylor polynomials of the policy function $Y_t^o = g_o(X_t, \varepsilon_t, \sigma)$ and $Y_t^n = g_n(X_t, \varepsilon_t, \sigma)$ of order s through perturbation methods.

2. Invert the Taylor polynomials using a Lagrange inversion formula to obtain $\varepsilon_t = \tilde{m}_{o,s}(X_t, Y_t^o, \sigma)$ and $Y_t^n = \tilde{m}_{n,s}(X_t, Y_t^o, \sigma)$.

3. Calculate the mode of $\log(\tilde{\pi}_{y,s}(Y_t^o|X_t))$, and the first M order derivatives at the mode.

4. Use the mode and first M order derivatives to construct the Laplace approximation.

The procedure is started at $t = 1$ with $X_1 = 0$, which assumes that the initial value is the deterministic steady state. Because Y_t contains X_{t+1}, for each t we can obtain X_{t+1} by using the observed values Y_t^o and the relationship $Y_t^n = \tilde{m}_{n,s}(X_t, Y_t^o, \sigma)$. Because the dimensions of ε_t and Y_t^o are the same, we do not need any Kalman Filter to calculate the likelihood. Note that when ε_t is normally distributed, using $s = 1$ gives the same likelihood as in the literature for linear DSGE models (e.g. Fernández-Villaverde (2010)).

Using the Newton-Raphson algorithm, we calculate the mode Y_L as the point that maximizes $\tilde{\pi}_{(y,s)}(Y_t^o|X_t)$ in (3.5). The Laplace approximation (3.4) is a polynomial-normal density with μ equal to Y_L and Ω equal to the inverse of minus the Hessian of $\log \tilde{\pi}_{(y,s)}(Y_t^o|X_t)$. The procedure is very fast because the derivatives are available in analytical form, and the mode for different values of t can be calculated in parallel, as we did in our code.

The following propositions provide the formulas for implementing the procedure for order $s = M = 2$, with the proofs available in Appendix A. Proposition 3.1 gives the Hessian and gradient of $\log \pi_y(Y_t^o|X_t)$ at any point Y_t^o as a function of the derivatives of $\hat{m}_o(X_t, Y_t^o, \sigma)$ when ε_t is normally distributed.

Proposition 3.2 gives closed expressions for the gradient and Hessian of $\log \tilde{\pi}_{(y,2)}(Y_t^o|X_t)$. Higher derivatives can be obtained using similar rules of matrix calculus, and are provided in the authors' personal website for the 3rd order.

Although there are numerous papers that provide general formulas for the Lagrange inversion (e.g. Apostol (2000), Johnson (2002)), Proposition 3.3 explains how to obtain it with the Kronecker product notation used by Dynare (2021). Leon-Gonzalez and Baiaman kyzy (2024) generalize this to the case $s > 2$.

Proposition 3.1. *Define J as the $n_o \times n_o$ Jacobian of $\hat{m}_o(X_t, Y_t^o, \sigma)$*

$$J = \frac{\partial \hat{m}_o(X_t, Y_t^o, \sigma)}{\partial (Y_t^o)'} \tag{3.6}$$

Let $Y_t^o = (Y_{t,1}^o, ..., Y_{t,n_o}^o)'$ and define F_i as the $n_o \times n_o$ matrix:

$$F_i = \frac{\partial J}{\partial Y_{t,i}^o}, \ \ i = 1, ..., n_o \tag{3.7}$$

and C as a $1 \times n_o$ vector

$$C = \begin{pmatrix} tr(J^{-1}F_1) & ... & tr(J^{-1}F_{n_o}) \end{pmatrix} \tag{3.8}$$

where $tr(.)$ is the trace operator. The gradient of $\log(\pi_y(Y_t^o|X_t))$ with respect to Y_t^o is

$$\frac{\partial \log(\pi_y(Y_t^o|X_t))}{\partial (Y_t^o)'} = -(\hat{m}_o(X_t, Y_t^o, \sigma))'\Sigma^{-1}J + C. \tag{3.9}$$

Let A be a $n_o \times n_o$ matrix defined as:

$$A = (a_{ij}), \ \ \text{where } a_{ij} = tr(J^{-1}F_iJ^{-1}F_j) \tag{3.10}$$

and let V be a $n_o \times n_o$ matrix defined as:

$$V = \begin{pmatrix} V_1 & ... & V_{n_o} \end{pmatrix}, \text{ where } V_i = -F_i'\Sigma^{-1}(\hat{m}_o(X_t, Y_t^o, \sigma)). \tag{3.11}$$

Then the Hessian of $\log(\pi_y(Y_t^o|X_t))$ with respect to Y_t^o is

$$H = -J'\Sigma^{-1}J + V - A + F \tag{3.12}$$

where F is the $R \times R$ matrix, defined as
$F = (f_{ij})$, where $f_{i,j} = tr(J^{-1}F_{ij})$ and F_{ij} is the $n_o \times n_o$ matrix defined as:

$$F_{i,j} = \frac{\partial F_i}{\partial Y_{t,j}^o}$$

Proposition 3.2. *Assume that $\varepsilon_t = \tilde{m}_{o,2}(X_t, Y_t^o, \sigma)$ is given by*

$$\varepsilon_t = \tilde{G}_o^{0,0} + \tilde{G}_o^{1,0}Y_t^o + \tilde{G}_o^{0,1}X_t + \tilde{G}_o^{2,0}(Y_t^o \otimes Y_t^o) + \tilde{G}_o^{0,2}(X_t \otimes X_t) + \tilde{G}_o^{1,1}(Y_t^o \otimes X_t) \tag{3.13}$$

where $\tilde{G}_o^{0,0}, \ldots, \tilde{G}_o^{1,1}$ are comformable matrices.

Then the Jacobian is:

$$J = \frac{\partial \tilde{m}_{o,2}(X_t, Y_t^o, \sigma)}{\partial (Y_t^o)'} = \tilde{G}_o^{1,0} + 2\tilde{G}_o^{2,0}(I_{n_o} \otimes Y_t^o) + \tilde{G}_o^{1,1}(I_{n_o} \otimes X_t) \tag{3.14}$$

where I_{n_o} is the identity matrix of dimension n_o.

Let i_j denote the j^{th} column of the identity matrix, such that $I_{n_o} = (i_1, ..., i_{n_o})$. Define C as the $n_o \times n_o$ matrix whose j^{th} column is equal to $2tr(J^{-1}\tilde{G}_o^{2,0}(I_{n_o} \otimes i_j))$, such that:

$$C = 2(tr(J^{-1}\tilde{G}_o^{2,0}(I_{n_o} \otimes i_1)), ..., tr(J^{-1}\tilde{G}_o^{2,0}(I_{n_o} \otimes i_{n_o}))) \tag{3.15}$$

The gradient of $\log(\tilde{\pi}_{(y,2)}(Y_t^o|X_t))$ with respect to Y_t^o is:

$$\frac{\partial \log(\pi_y(Y_t^o|X_t))}{\partial (Y_t^o)'} = -(\tilde{m}_{o,2}(X_t, Y_t^o, \sigma))'\Sigma^{-1}J + C \tag{3.16}$$

Let A be a $n_o \times n_o$ matrix defined as:

$$A = (a_{ij}), \text{ where } a_{ij} = 4tr(J^{-1}\tilde{G}_o^{2,0}(I_{n_o} \otimes i_i)J^{-1}\tilde{G}_o^{2,0}(I_{n_o} \otimes i_j)) \tag{3.17}$$

Let V be a $n_o \times n_o$ matrix defined as:

$$V = \begin{pmatrix} V_1 & \dots & V_{n_o} \end{pmatrix}, \text{ where } V_i = -(2\tilde{G}_o^{2,0}(I_{n_o} \otimes i_i))'\Sigma^{-1}(\tilde{m}_{o,2}(X_t, Y_t^o, \sigma)) \tag{3.18}$$

The Hessian of $\log(\tilde{\pi}_{(y,2)}(Y_t^o|X_t))$ *with respect to* Y_t^o *is:*

$$H = -J'\Sigma^{-1}J + V - A \tag{3.19}$$

Proposition 3.3. *Let the Taylor Polynomials* $\varepsilon_t = \tilde{m}_{o,2}(X_t, Y_t^o, \sigma)$ *and* $Y_t^n = \tilde{m}_{n,2}(X_t, Y_t^o, \sigma)$ *be given by:*

$$\begin{aligned} \varepsilon_t &= \tilde{G}_o^{0,0} + \tilde{G}_o^{1,0}Y_t^o + \tilde{G}_o^{0,1}X_t + \tilde{G}_o^{2,0}(Y_t^o \otimes Y_t^o) + \tilde{G}_o^{0,2}(X_t \otimes X_t) + \tilde{G}_o^{1,1}(Y_t^o \otimes X_t) \\ Y_t^n &= \tilde{G}_n^{0,0} + \tilde{G}_n^{1,0}Y_t^o + \tilde{G}_n^{0,1}X_t + \tilde{G}_n^{2,0}(Y_t^o \otimes Y_t^o) + \tilde{G}_n^{0,2}(X_t \otimes X_t) + \tilde{G}_n^{1,1}(Y_t^o \otimes X_t) \end{aligned} \tag{3.20}$$

where $\tilde{G}_o^{0,0}, \dots, \tilde{G}_o^{1,1}$, *and* $\tilde{G}_n^{0,0}, \dots, \tilde{G}_n^{1,1}$ *are comformable matrices. The* $\tilde{G}$ *matrices in (3.20) can be obtained from the* G *matrices in (2.4) as follows.*

$$\begin{cases} \tilde{G}_o^{1,0} = \Big(G_o^{1,0} - 2G_o^{2,0}\Big(G_o^{1,0} \otimes G_o^{1,0} + 2(G_o^{2,0} \otimes G_o^{0,0})\Big)^{-1}(G_o^{0,0} \otimes G_o^{1,0})\Big)^{-1} \\ \\ \tilde{G}_n^{1,0} = \Big(G_n^{1,0} - 2G_n^{2,0}\Big(G_o^{1,0} \otimes G_o^{1,0} + 2(G_o^{2,0} \otimes G_o^{0,0})\Big)^{-1}(G_o{}^{0,0} \otimes G_o{}^{1,0})\Big)\tilde{G}_o^{1,0} \end{cases} \tag{3.21}$$

$$\begin{cases} \tilde{G}_o^{2,0} = -\tilde{G}_o^{1,0}G_o^{2,0}\Big(G_o^{1,0} \otimes G_o^{1,0} + 2(G_o^{2,0} \otimes G_o^{0,0})\Big)^{-1} \\ \\ \tilde{G}_n^{2,0} = \Big(G_n{}^{2,0} - \tilde{G}_n^{1,0}G_o{}^{2,0}\Big)\Big(G_o^{1,0} \otimes G_o^{1,0} + 2(G_o^{2,0} \otimes G_o^{0,0})\Big)^{-1} \end{cases} \tag{3.22}$$

$$
\begin{cases}
\tilde{G}_o^{1,1} = -\Big(\tilde{G}_o^{1,0} G_o^{1,1} + 2\tilde{G}_o^{2,0}(G_o^{1,0} \otimes G_o^{0,1}) + 2\tilde{G}_o^{2,0}(G_o^{1,1} \otimes {G_o}^{0,0})\Big)\Big((G_o^{1,0})^{-1} \otimes I_{n_K}\Big) \\
\\
\tilde{G}_n^{1,1} = \Big(G_n^{1,1} - \Big(\tilde{G}_n^{1,0} G_o^{2,0} + 2\tilde{G}_n^{2,0}({G_o}^{1,0} \otimes {G_o}^{0,1})\Big) + 2\tilde{G}_n^{2,0}(G_o^{1,1} \otimes {G_o}^{0,0})\Big)\Big(({G_o}^{1,0})^{-1} \otimes I_{n_K}\Big)
\end{cases}
\tag{3.23}
$$

$$
\begin{cases}
\tilde{G}_o^{0,2} = -\Big(\tilde{G}_o^{1,0} G_o^{0,2} + \tilde{G}_o^{2,0}(G_o^{0,1} \otimes G_o^{0,1}) + \tilde{G}_o^{1,1}(G_o^{0,1} \otimes I_{n_K})\Big) - 2\tilde{G}_o^{2,0}\Big({G_o}^{0,0} \otimes {G_o}^{0,2}\Big) \\
\\
\tilde{G}_n^{0,2} = G_n^{0,2} - \Big(\tilde{G}_n^{1,0} G_o^{0,2} + \tilde{G}_n^{2,0}(G_o^{0,1} \otimes G_o^{0,1}) + \tilde{G}_n^{1,1}(G_o^{0,1} \otimes I_{n_K})\Big) - 2\tilde{G}_n^{2,0}\Big(G_o^{0,0} \otimes G_o^{0,2}\Big)
\end{cases}
\tag{3.24}
$$

$$
\begin{cases}
\tilde{G}_o^{0,1} = -\Big(\tilde{G}_o^{1,0} G_o^{0,1} + \tilde{G}_o^{1,1}(G_o^{0,0} \otimes I_{n_K})\Big) - 2\tilde{G}_o^{2,0}\Big(G_o^{0,0} \otimes G_o^{0,1}\Big) \\
\\
\tilde{G}_n^{0,1} = G_n^{0,1} - \Big(\tilde{G}_n^{1,0} G_o^{0,1} + \tilde{G}_n^{1,1}(G_o^{0,0} \otimes I_{n_K})\Big) - 2\tilde{G}_n^{2,0}\Big(G_o^{0,0} \otimes G_o^{0,1}\Big)
\end{cases}
\tag{3.25}
$$

$$
\begin{cases}
\tilde{G}_o^{0,0} = -\Big(\tilde{G}_o^{1,0} G_o^{0,0} + \tilde{G}_o^{2,0}(G_o^{0,0} \otimes G_o^{0,0})\Big) \\
\\
\tilde{G}_n^{0,0} = G_n^{0,0} - \Big(\tilde{G}_n^{1,0} G_o^{0,0} + \tilde{G}_n^{2,0}(G_o^{0,0} \otimes G_o^{0,0})\Big)
\end{cases}
\tag{3.26}
$$

4 Simulation from the Laplace Based Solution to the DSGE Model

Once the model has been solved by perturbation methods, equation (2.4) can be used to simulate directly values for Y_t. For a given value of X_1 this can be done by repeating the following two steps for $t = 1, ..., T$:

1. Simulate ε_t from the appropriate distribution and use equation (2.4) to obtain Y_t.
2. Obtain X_{t+1} as the appropriate subvector of Y_t.

We use this approach to obtain the generalized Impulse Response Functions (IRFs) presented in Section 5 (see e.g. Dynare (2021) for an explanation of how to use simulation to construct IRFs).

However, it is also possible to simulate Y_t using the likelihood of the Laplace based solution that we have proposed in Section 3. Specifically, for a given initial value of X_1 this can be done by repeating the following two steps for $t = 1, ..., T$:

1. Simulate Y_t^o using the Laplace approximated density $\hat{\pi}_{(y,M)}(Y_t^o|X_t)$.
2. Obtain Y_t^n using the Lagrange inverse $Y_t^n = \tilde{m}_{n,s}(X_t, Y_t^o, \sigma)$. Obtain X_{t+1} as the appropriate subvector of Y_t.

Note that perturbation methods give an approximation of the policy functions around the steady state, and that the quality of the approximation deteriorates as we get further from the steady state. For this reason, we cannot expect the perturbation methods to be informative about the tails of the distribution of Y_t. Instead, we can expect that several distributions will be consistent with the local properties of the true distribution around the steady state. The Laplace based solution of the model uses the local derivatives to construct a density which is consistent with the local properties around the steady state, and yet can be calculated easily numerically. In contrast, as argued in previous sections, the exact likelihood for the solution obtained using only perturbation methods cannot be computed, except in limited cases.

However, if perturbation methods are accurate in approximating the true policy functions, we should expect both solutions to give similar results. We can evaluate this by comparing the IRFs obtained from both approaches. To obtain IRFs in the Laplace based solution, we should introduce intervention dummies i_t containing the impulses. Hence, the structural errors of this economy become $\varepsilon_t + i_t$, which we introduce in the Lagrange inverse polynomials $\varepsilon_t + i_t = \tilde{m}_{o,2}(X_t, Y_t^o, \sigma)$ and $Y_t^n = \tilde{m}_{n,2}(X_t, Y_t^o, \sigma)$ that were presented in equation (3.20) as follows:

$$\begin{aligned}\varepsilon_t &= \tilde{G}_o^{0,0} - i_t + \tilde{G}_o^{1,0} Y_t^o + \tilde{G}_o^{0,1} X_t + \tilde{G}_o^{2,0}(Y_t^o \otimes Y_t^o) + \tilde{G}_o^{0,2}(X_t \otimes X_t) + \tilde{G}_o^{1,1}(Y_t^o \otimes X_t) \\ Y_t^n &= \tilde{G}_n^{0,0} + \tilde{G}_n^{1,0} Y_t^o + \tilde{G}_n^{0,1} X_t + \tilde{G}_n^{2,0}(Y_t^o \otimes Y_t^o) + \tilde{G}_n^{0,2}(X_t \otimes X_t) + \tilde{G}_n^{1,1}(Y_t^o \otimes X_t)\end{aligned} \tag{4.1}$$

These equations are the same as in (3.20) except that we now write $(\varepsilon_t + i_t)$ instead of (ε_t) or equivalently $(\tilde{G}_o^{0,0} - i_t)$ instead of $(\tilde{G}_o^{0,0})$. The simulation then can be carried out as explained above but using (4.1) instead of (3.20), and choosing the vector i_t according to the IRF that needs to be calculated.

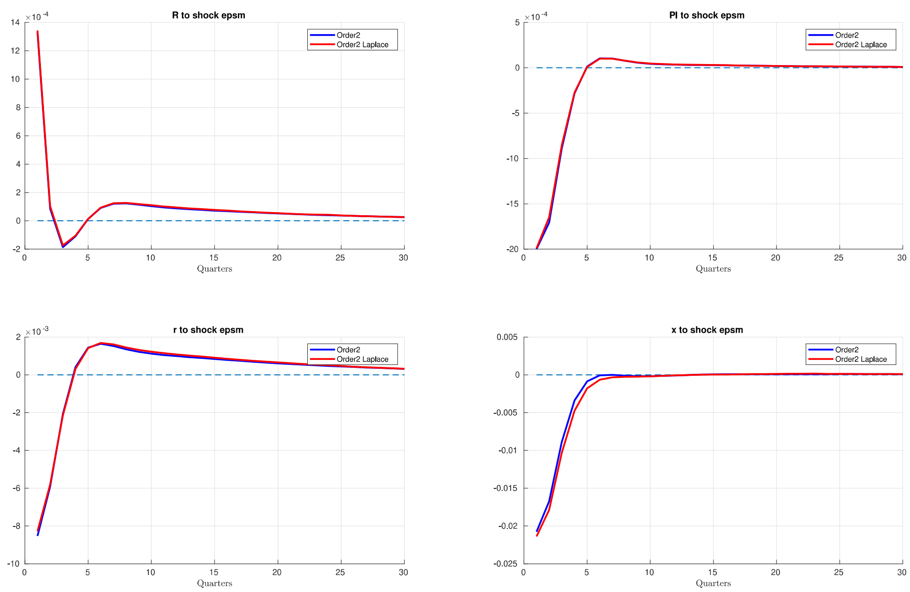


Figure 1: IRFs to a Monetary Policy Shock: Second-Order Perturbation Methods versus Laplace Based Solution

Figure 1 shows the IRFs to a monetary policy shock obtained with these two approaches for the empirical analysis in Section 5. We find that the IRFs are very similar and mostly overlap with each other[2].

[2]The figure shows the responses of inflation, nominal interest rate, investment and rental rate of capital. The responses of other variables are also similar, and can be found in Baiaman kyzy (2023).

5 Empirical Analysis

5.1 Model and Data Description

We estimate linear and nonlinear variants of the neoclassical growth DSGE model of Fernández-Villaverde (2010) using US data from 1959 Q1 to 2019 Q4 (244 observations). We used the same 5 variables as in Fernández-Villaverde (2010), updated to a longer sample: 1) the relative price of investment with respect to the price of consumption, 2) real output per capita growth, 3) real wages per capita growth, 4) the consumer price index growth and 5) the federal funds rate. Table B3 in Appendix B describes the data sources and the transformations to obtain the observed variables in the model (Y_t^o). For the nonlinear models we used a second order approximation ($s = M = 2$) in the logs of the variables.

The model has 5 normally distributed structural shocks: a preference shock $\varepsilon_{d,t}$, a labor disutility shock $\varepsilon_{\varphi,t}$, an investment specific technology shock $\varepsilon_{\mu_I,t}$, a neutral technology shock $\varepsilon_{A,t}$ and a monetary policy shock m_t. Their standard deviations are estimated and are denoted as $\exp \sigma_d$, $\exp \sigma_\varphi$, $\exp \sigma_\mu$, $\exp \sigma_A$ and $\exp \sigma_m$, respectively. The model is defined by 30 equations (Table B1) and has 30 variables (Table B2). The structural errors appear in equations (6.24)- (6.27) and (6.15) in Table B1. We estimate 25 unknown parameters in the linear and homoscedastic nonlinear models. We follow Fernández-Villaverde (2010) in the specification of the priors (Tables B4 and B5) and in the calibration of some parameters: $\varepsilon = 10$, $\eta = 10$ and $\phi = 0$. However, we differ in leaving unrestricted two parameters: δ and γ_2.

The basic structure of this neoclassical growth model is as follows. There is a representative household which consumes, saves, holds money, supplies labor, and sets its own wages subject to a demand curve and Calvo's pricing. The final output is produced by a final good firm, which uses as inputs a continuum of intermediate goods manufactured by monopolistic competitors. The intermediate good producers rent capital and labor to produce their good. They face the constraint that they can only change prices following a Calvo's rule. Finally, there is a monetary authority that fixes the one-period nominal interest rate through open market operations with public debt.

One of the limitations of the linear DSGE models is that they cannot handle heteroscedastic structural errors, because the heteroscedasticity disappears from the model with the linear approximation. However, heteroscedasticity is important to reflect the changing uncertainty in the economy, and it also greatly improves the empirical performance of econometric models. We therefore consider two heteroscedastic versions of the nonlinear model. The first one allows the variances of the structural errors to change in the fashion of a GARCH model. The shocks continue to have the same unconditional variance (e.g. $\exp(2\sigma_d)$), but they are multiplied by a time-varying process with an expected value of one:

$$\log \tilde{d}_t = \rho^d \log d_{t-1} + \sqrt{\tilde{\sigma}_{d,t}} \varepsilon_{d,t}$$

$$\tilde{\sigma}_{d,t} = \rho_1^d \tilde{\sigma}_{d,t-1} + \rho_2^d \tilde{\sigma}_{d,t-1} \tfrac{(\varepsilon_{d,t-1})^2}{\exp(2\sigma_d)} + (1 - \rho_1^d - \rho_2^d)$$

$$\log \varphi_t = \rho^\varphi \log \varphi_{t-1} + \sqrt{\tilde{\sigma}_{\varphi,t}} \varepsilon_{\varphi,t}$$

$$\tilde{\sigma}_{\varphi,t} = \rho_1^\varphi \tilde{\sigma}_{\varphi,t-1} + \rho_2^\varphi \tilde{\sigma}_{\varphi,t-1} \tfrac{(\varepsilon_{\varphi,t-1})^2}{\exp(2\sigma_\varphi)} + (1 - \rho_1^\varphi - \rho_2^\varphi)$$

$$\log \mu_{I,t} = \Lambda_\mu + \sqrt{\tilde{\sigma}_{\mu,t}} \varepsilon_{\mu_I,t}$$

$$\tilde{\sigma}_{\mu,t} = \rho_1^\mu \tilde{\sigma}_{\mu,t-1} + \rho_2^\mu \tilde{\sigma}_{\mu,t-1} \tfrac{(\varepsilon_{\mu_I,t-1})^2}{\exp(2\sigma_\mu)} + (1 - \rho_1^\mu - \rho_2^\mu)$$

$$\log \mu_{A,t} = \Lambda_A + \sqrt{\tilde{\sigma}_{A,t}} \varepsilon_{A,t}$$

$$\tilde{\sigma}_{A,t} = \rho_1^A \tilde{\sigma}_{A,t-1} + \rho_2^A \tilde{\sigma}_{A,t-1} \tfrac{(\varepsilon_{A,t-1})^2}{\exp(2\sigma_A)} + (1 - \rho_1^A - \rho_2^A)$$

$$\tfrac{R_t}{R} = \left(\tfrac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\tfrac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{\frac{\tilde{y}_t^d}{\tilde{y}_{t-1}^d} \frac{z_t}{z_{t-1}}}{\Lambda_{y^d}} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp\left(\sqrt{\tilde{\sigma}_{m,t}} m_t\right)$$

$$\tilde{\sigma}_{m,t} = \rho_1^m \tilde{\sigma}_{m,t-1} + \rho_2^m \tilde{\sigma}_{m,t-1} \tfrac{(\varepsilon_{m,t-1})^2}{\exp(2\sigma_m)} + (1 - \rho_1^m - \rho_2^m)$$

where the last two equations define monetary policy using a Taylor rule. We assume that $\rho_1^i > 0, \rho_2^i > 0$ and that $(\rho_1^i + \rho_2^i) < 1$, for $i = d, \varphi, \mu, A, m$. Under these restrictions the time-varying variances ($\tilde{\sigma}_{d,t}$, $\tilde{\sigma}_{\varphi,t}$, $\tilde{\sigma}_{\mu,t}$, $\tilde{\sigma}_{A,t}$, $\tilde{\sigma}_{m,t}$) have expected values equal to one, so that the long-run variances are ($\exp 2\sigma_d$, $\exp 2\sigma_\varphi$, $\exp 2\sigma_\mu$, $\exp 2\sigma_A$, $\exp 2\sigma_m$). The GARCH version of the model has 10 extra parameters: ρ_1^d, ρ_2^d, ρ_1^φ, ρ_2^φ, ρ_1^μ, ρ_2^μ, ρ_1^A, ρ_2^A, ρ_1^m, ρ_2^m. We specify independent beta priors for each of these parameters. The prior mean and standard deviation for ρ_1^i are 0.7 and 0.046, respectively, for $i = d, \varphi, \mu, A, m$. The prior mean and standard deviation for ρ_2^i are 0.2 and 0.12, respectively, for $i = d, \varphi, \mu, A, m$.

We also consider a restricted GARCH model in which the variances evolve according to a common multiplicative factor $\widetilde{\sigma}_t$. Here we assume that for every t this restriction holds: $\widetilde{\sigma}_t$=$\tilde{\sigma}_{d,t}$= $\tilde{\sigma}_{\varphi,t}$= $\tilde{\sigma}_{\mu,t}$= $\tilde{\sigma}_{A,t}$= $\tilde{\sigma}_{m,t}$. The common factor $\widetilde{\sigma}_t$ responds to past values of the structural shocks as in a GARCH model:

$$\widetilde{\sigma}_t = \rho_1 \widetilde{\sigma}_{t-1} + \rho_2 \widetilde{\sigma}_{t-1} \tfrac{(\widetilde{\varepsilon}_{t-1})^2}{var(\widetilde{\varepsilon}_t)} + (1 - \rho_1 - \rho_2)$$

$$\widetilde{\varepsilon}_t = \delta_d \tfrac{\varepsilon_{d,t}}{\exp(\sigma_d)} + \delta_\varphi \tfrac{\varepsilon_{\varphi,t}}{\exp(\sigma_\varphi)} + \delta_\mu \tfrac{\varepsilon_{\mu_I,t}}{\exp(\sigma_\mu)} + \delta_A \tfrac{\varepsilon_{A,t}}{\exp(\sigma_A)} + \delta_m \tfrac{\varepsilon_{m,t}}{\exp(\sigma_m)}$$

$$var(\widetilde{\varepsilon}_t) = \delta_d^2 + \delta_\varphi^2 + \delta_\mu^2 + \delta_A^2 + \delta_m^2$$

In this model all structural shocks contribute to the time variation of the common factor. We can measure the relative contributions of the shocks to such time variation by the following proportions:

$$p_d = \frac{\delta_d^2}{var(\widetilde{\varepsilon}_t)},\ p_\varphi = \frac{\delta_\varphi^2}{var(\widetilde{\varepsilon}_t)},\ p_\mu = \frac{\delta_\mu^2}{var(\widetilde{\varepsilon}_t)},\ p_A = \frac{\delta_A^2}{var(\widetilde{\varepsilon}_t)},\ p_m = \frac{\delta_m^2}{var(\widetilde{\varepsilon}_t)},$$

where $p_d + p_\varphi + p_\mu + p_A + p_m = 1$. For example, p_m is the proportion of the time-variation in uncertainty driven by the monetary policy shock.

When it comes to estimating the common factor GARCH model, we have to realize that we have to normalize the vector $\delta = (\delta_d, \delta_\varphi, \delta_\mu, \delta_A, \delta_m)$ because it is not identified. We normalize it by the restriction $\delta_d^2 + \delta_\varphi^2 + \delta_\mu^2 + \delta_A^2 + \delta_m^2 = 1$, such that $var(\widetilde{\varepsilon}_t) = 1$. Regarding the prior, we specify a beta prior for ρ_1, with mean 0.7 and standard deviation 0.046. We then define $\widetilde{\delta} = \sqrt{\rho_2}\delta$ such that $\rho_2 = \widetilde{\delta}'\widetilde{\delta}$, and specify a normal prior for $\widetilde{\delta}$ with 0 mean and var-cov matrix equal to $(0.2/5)I_5$, where I_5 is the identity matrix. This implies that the prior for ρ_2 is a chi-squared distribution with 5 degrees of freedom and mean equal to 0.2. The implied prior for the normalized vector δ is a uniform.

We, therefore, estimate and compare four models:

- M_1: Log-linearized model with homoscedastic shocks.
- M_2: Model resulting from second-order approximation with homoscedastic shocks.
- $M_{2,G}$: Model resulting from second-order approximation with unrestricted GARCH in the shocks.

- $M_{2,fG}$: Model resulting from second-order approximation with a common factor in the GARCH processes.

5.2 Results and Discussion

Models are evaluated with the marginal likelihood (e.g., Koop (2003), p. 4), predictive likelihood (Geweke and Amisano (2010)), and posterior probabilities that use the predictive likelihood as weights. The marginal likelihood was calculated following Gelfand and Dey (1994) and Geweke (1999), and the predictive likelihood corresponds to observations from 45 to 244. We used a block Metropolis-Hasting algorithm with a random walk proposal to obtain draws from the posterior distribution (e.g., Koop (2003), p. 97). The average computation time for estimation of models is 2–3 hours with 300000 iterations[3].

Table 1 shows that all nonlinear models (M_2, $M_{2,G}$, and $M_{2,fG}$) are better than the linear one (M_1) in terms of marginal likelihoods and predictive likelihoods, with the nonlinear factor GARCH model $M_{2,fG}$ being by far the best performing model with a posterior probability of one. Therefore, we can conclude that the nonlinear solution fits the data better than a linear approximation, and that volatilities follow a common pattern over time.

Regarding the common factor GARCH model $M_{2,fG}$, Table 2 shows the posterior estimates for the proportions of uncertainty caused by each structural shock. We find that the monetary policy shock causes 95% of the time-variation in volatilities, whereas the investment specific technology shock causes 4%. The posterior estimates and 95% credible intervals in the four models for all parameters can be found in Appendix B in Tables B4, B5, B6, B7 and B8.

Figure 2 shows the generalized impulse response functions to a positive monetary policy shock, which represent log deviations from the steady-state following a shock of one standard deviation. As expected, a positive monetary shock has a contractionary impact on aggregate output, consumption, real wages, and investment in all four models. However, model M_2 results in IRFs that are somewhat different from those in other models. This is because the posterior estimates for some parameters (for example h, γ, κ, and χ) are different, causing different steady-state values for consumption, investment, output, and wages (Table B9), which in turn induces slightly different IRFs. If we used

[3]Using a computer with processor Intel(R) Core (TM) i7-10700, CPU 2.90GHz, and RAM 16.0 GB.

the same values for the parameters, models M_1 and M_2 would provide similar IRFs (as shown in Figure B7).

Figure 3 plots the likelihood for each observation, which measures the probability of observing the data for each quarter based on the estimated parameters. The sample period captures several main recessions, mainly the economic downturn in the US followed by the oil price shock of the 1970s or Early 80s Recession; and the Great Recession of 2008 to 2009. During both periods, the log-likelihood for models with heteroscedastic shocks decreases less than for the linear and homoscedastic quadratic models. The reason for this might be that in heteroscedastic models the conditional variance in periods of crisis increases, and therefore the likelihood decreases less when there is a large shock. Figure B8 in Appendix B shows the cumulative likelihood function over the sample, showing that $M_{2,fG}$ outperforms other models and better captures the characteristics of the data.

Appendix B shows the fitted and actual values (Figures B1-B4) as well as posterior estimates of latent processes such as productivity growth and marginal cost (Figures B5 and B6).

Model	**Nº of parameters**	**Log Marg. Likelihood**	**Predictive Likelihood**	**Pr(M|Y)**
M_1	25	4596.30	3782.58	0
M_2	25	4602.10	3794.59	0
$M_{2,G}$	35	4639.30	3792.19	0
$M_{2,fG}$	31	4711.60	3859.99	1

Table 1: Model Performance Measures for M_1, M_2, $M_{2,G}$ and $M_{2,fG}$. The log marginal likelihood was calculated by importance sampling (Geweke (1999)). The predictive likelihood is calculated for observations 45 to 244. $Pr(M|Y)$ is the posterior probability of model M using the predictive likelihood.

	p_d	p_φ	p_μ	p_A	p_m
Estimate	0.0039	0.0001	0.0434	0.0012	0.9514

Table 2: Posterior Means for Proportions in Model $M_{2,fG}$

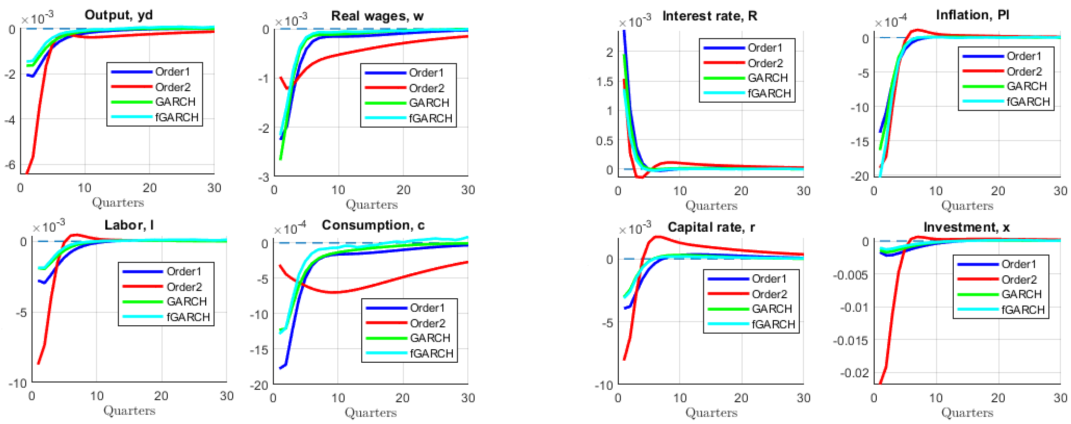


Figure 2: IRFs to a Positive Monetary Policy Shock. The IRFs are calculated using the posterior mode of the parameters.

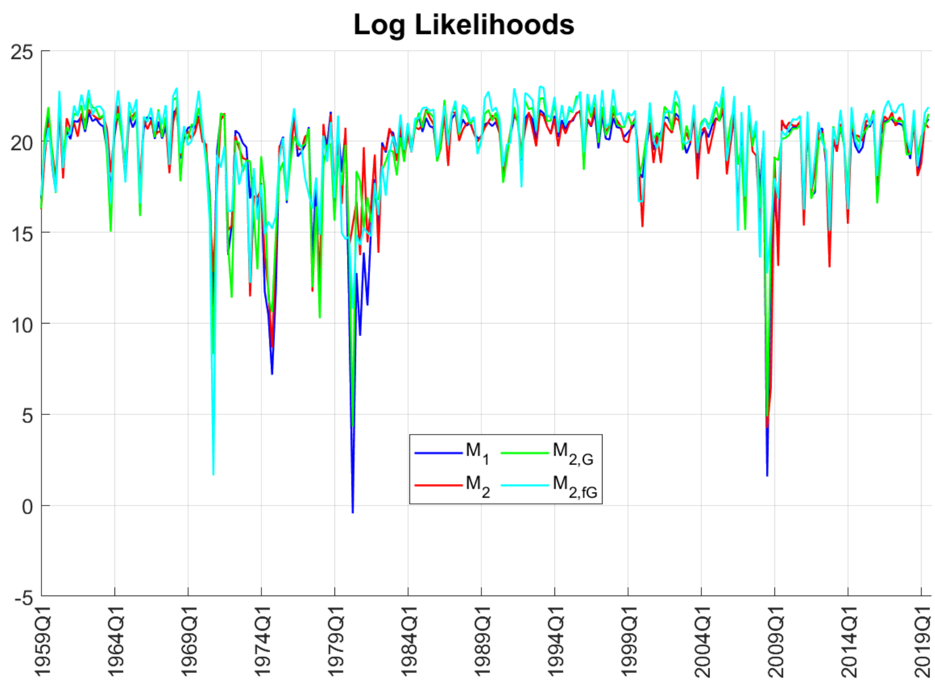


Figure 3: Log Likelihood for Models M_1, M_2, $M_{2,G}$ and $M_{2,fG}$. The likelihood is evaluated at the posterior mode of the parameters.

6 Conclusions

This paper developed a new methodology for nonlinear DSGE models that allows the likelihood-based estimation of some important models that cannot be estimated with the current likelihood methods available in the literature. In particular, this method neither requires the number of shocks to be greater than the number of observed variables, nor that some shocks enter linearly in the model. In addition, the method allows for the presence of unobserved non-stochastic state variables, such as capital.

This method implicitly uses an invertible approximation of the policy function, according to which the shocks can be recovered uniquely from some of the control variables. It then uses the local inverse of the policy function to provide a Laplace based solution of the model. The likelihood of this solution can be calculated in closed form, which greatly speeds the calculation and allows for the estimation of larger models. In contrast with previous methods, our novel methodology does neither require simulation to evaluate the likelihood, nor solving systems of polynomial equations.

In the empirical analysis we used our novel method to estimate linear and nonlinear variants of the well-known neoclassical growth model of Fernández-Villaverde (2010) using US data from 1959 Q1 to 2019 Q4. Although the linear version can only handle homoscedastic structural errors, among the nonlinear variants we considered both homoscedastic and heteroscedastic structural errors. We found that all nonlinear variants performed better than the linear model in terms of the marginal and predictive likelihoods. A novel common factor GARCH model had by far the best performance and allowed us to estimate that 95% of the economic uncertainty was caused by the monetary policy shock. Although the IRFs in this model were similar to the ones obtained with the linear model, the heteroscedastic models had much higher likelihood values in periods of economic turbulence, thanks to their better representation of uncertainty.

References

Altug, S. (1989). Time-to-Build and Aggregate Fluctuations: Some New Evidence. *International Economic Review*, 30, 889–920.

Amisano, G., & Tristani, O. (2011). Exact Likelihood Computation for Nonlinear DSGE Models with Heteroskedastic Innovations. *Journal of Economic Dynamics & Control*, 35, 2167–2185. https://doi.org/10.1016/j.jedc.2011.08.003

Apostol, T.M. (2000). Calculating Higher Derivatives of Inverses. *The American Mathematical Monthly*, 107, 738–741. https://doi.org/10.2307/2695472

Baiaman kyzy, E. (2023) *Nonlinear Analysis and Estimation of Dynamic Stochastic General Equilibrium Models*. Doctoral Thesis, National Graduate Institute for Policy Studies, No 277. https://doi.org/10.24545/0002000044

Billingsley, P. (1999). *Convergence of Probability Measures*. Jonh Willey Sons, Inc.

Blanchard, O. J., & Kahn, C. M. (1985). The Solution of Linear Difference Models Under Rational Expectations. *Econometrica*, 48 (5), 1305–1311. https://doi.org/10.2307/1912186

Christiano, L. J. (1990). Solving the Stochastic Growth Model by Linear-Quadratic Approximation and by Value-Function Iteration. *Journal of Business & Economic Statistics*, 8, 23–26.

Christiano, L. J., & Eichenbaum, M. (1992). Current Business Cycle Theories and Aggregate Labor Market Fluctuations. *American Economic Review*, 82, 430–450.

Collard, F., & Juillard, M. (2001). Accuracy of Stochastic Perturbation Methods: The Case of Asset Pricing Models. *Journal of Economic Dynamics & Control*, (25), 979–999. https://doi.org/10.1016/S0165-1889(00)00064-6

DeJong, D. N., Ingram, B. F., & Whiteman, C. H. (2000). A Bayesian Approach to Dynamic Macroeconomics. *Journal of Econometrics*, 98, 203–223.

Dynare. (2021). *The Dynare Reference Manual*. Version 4.6.4. Dynare team.

Fernández-Villaverde, J. (2010). The Econometrics of DSGE Models. *Journal of the Spanish Economic Association*, (1), 3–49. https://doi.org/10.1007/s13209-009-0014-7

Fernández-Villaverde, J., & Rubio-Ramírez, J. F. (2007). Estimating Macroeconomic Models: A Likelihood Approach. *Review of Economic Studies*, 74, 1059–1087. http://www.jstor.org/stable/4626173

Fleming, W. H. (1971). Stochastic Control for Small Noise Intensities. *SIAM Journal of Control*, 9, 473–517. https://doi.org/10.1137/0309035

Galewski, M. (2016). A Note on a Global Invertibility of Mappings on Rn. *Quaestiones Mathematicae*, 39 (5), 683–688. https://doi.org/10.2989/16073606.2015.1121933

Gelfand, A., & Dey, D. (1994). Bayesian Model Choice: Asymptotics and Exact Calculations. *Journal of the Royal Statistical Society Series B*, 56, 501–514.

Geweke, J. (1999). Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication (with discussion and rejoinder). *Econometric Reviews*, 18, 1–126.

Geweke, J., & Amisano, G. (2010). Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns. *International Journal of Forecasting*, 26 (2), 216–230. https://doi.org/10.1016/j.ijforecast.2009.10.007

Gordon, G., & Qiu, S. (2018). A Divide and Conquer Algorithm for Exploiting Policy Function Monotonicity. *Quantitative Economics*, 9 (2), 521–540. https://doi.org/10.3982/QE640

Hansen, L. P., & Sargent, T. J. (1980). Formulating and Estimating Dynamic Linear Rational Expectations Models. *Journal of Economic Dynamic and Control*, 2, 7–46.

Hansen, L. P., & Singleton, K. (1983). Stochastic Consumption, Risk Aversion and Temporal Behavior of Asset Returns. *Journal of Political Economy*, 91, 249–265.

Ishikawa, I., Teshima, T., Tojo, K., Oono, K., Ikeda, M., & Sugiyama, M. (2022). Universal Approximation Property of Invertible Neural Networks. *arXiv:2204.07415v1 [cs.LG]*. https://doi.org/10.48550/arXiv.2204.07415

Johnson, W. P. (2002). Combinatorics of Higher Derivatives of Inverses. *The American Mathematical Monthly*, 109 (3), 273-277, shttps://doi.org/10.1080/00029890.2002.11919861

Judd, K. L. (1998). *Numerical Methods in Economics*. MIT Press, Cambridge.

Klein, P. (2000). Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics & Control*, (24), 1405–1423. https://doi.org/10.1016/S0165-1889(99)00045-7

Kollmann, R. (2017). Tractable Likelihood-Based Estimation of Non-Linear DSGE Models. *Economic Letters*, 161, 90-92.

Leon-Gonzalez, R. & Baiaman kyzy, E. (2024). Lagrange Inversion of Taylor Polynomials Using Kronecker Product Notation. *Manuscript*, https://sites.google.com/view/robertoleongonzalez.

Lipoth, J., Papalexiou, S. M., & Spiteri, R. J. (2022). A New Very Simply Explicitly Invertible Approximation for the Standard Normal Cumulative Distribution Function. *AIMS Mathematics*, 7 (7), 11635–11646. https://doi.org/10.3934/math.2022648

Koop, G. (2003). *Bayesian Econometrics*. Wiley-Interscience.

Otrok, C. (2001). On Measuring the Welfare Cost of Business Cycles. *Journal of Monetary Economics*, 47, 61–92.

Schmitt-Grohé, S. & Uribe, M. (2004). Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function. *Journal of Economic Dynamics & Control*, (28), 755–775. https://doi.org/10.1016/S0165-1889(03)00043-5

Schorfheide, F. (2000). Loss Function-Based Evaluation of DSGE Models. *Journal of Applied Econometrics*, 15 (6), 645–670. https://doi.org/10.1002/jae.582

Sims, C. A. (2000). *Second Order Accurate Solution of Discrete Time Dynamic Equilibrium Models*. Manuscript, Princeton University, Princeton. https://doi.org/10.2139/ssrn.484443

Sims, C. A. (2002). Solving Linear Rational Expectations Models. *Computational Economics*, 20, 1–20. https://doi.org/10.1023/A:1020517101123

Skoulakis, G. (2019). Simulating from Polynomial-Normal Distributions. *Communications in Statistics - Simulation and Computation*, 48, 472-477. https://doi.org/10.1080/03610918.2017.1385814

Teshima, T., Ishikawa, I., Tojo, K., Oono, K., Ikeda, M., & Sugiyama, M. (2020). Coupling-Based Invertible Neural Networks Are Universal Diffeomorphism Approximators. *Advances in Neural Information Processing Systems*, (33), 3362–3373.

Appendix A: Proofs of Propositions 3.1, 3.2 and 3.3

Proof of Proposition 3.1

Proof. Let $\pi_o = \log(\pi_y(Y_t^o|X_t))$. From (3.3), assuming normality for ε_t we have that:

$$\pi_o = \log(\pi_o(Y_t^o|X_t)) = -\frac{1}{2}\varepsilon_t^{'}\Sigma^{-1}\varepsilon_t + \log|J| - \frac{1}{2}\log|\Sigma| - \frac{n_o}{2}\log(2\pi)$$

Using matrix differential calculus, the derivatives of a determinant can be calculated (e.g Magnus and Neudecker (1999)), such that a differential of π_o can be written as:

$$\partial\pi_o = -\varepsilon_t^{'}\Sigma^{-1}\partial\varepsilon_t + tr(J^{-1}\partial J)$$

where ∂J is the differential of J and can be written as:

$$\partial J = \sum_{i=1}^{n_o}\frac{\partial J}{\partial Y_{t,i}^o}\partial Y_{t,i}^o = \sum_{i=1}^{n_o} F_i \partial Y_{t,i}^o$$

Therefore, we can write:

$$tr(J^{-1}\partial J) = tr(J^{-1}\sum_{i=1}^{n_o} F_i\partial Y_{t,i}^o) = \sum_{i=1}^{n_o} tr(J^{-1}F_i)\partial Y_{t,i}^o = C\partial Y_t^o$$

where: $\partial Y_t^o = (\partial Y_{t,1}^o, Y_{t,2}^o, \ldots Y_{t,n_o}^o)^{'}$.
Therefore, $\partial\pi_o$ can be written as:

$$\partial\pi_o = -\varepsilon_t^{'}\Sigma^{-1}\partial\varepsilon_t + C\partial Y_t^o$$

From the definition of Jacobian J, we have that $\partial\varepsilon_t = J\partial Y_t^o$. Therefore, we can write:

$$\partial\pi_o = -\varepsilon_t^{'}\Sigma^{-1}J\partial Y_t^o + C\partial Y_t^o$$

which gives the result that proves (3.9):

$$\frac{\partial\pi_o}{\partial(Y_t^o)^{'}} = -\varepsilon_t^{'}\Sigma^{-1}J + C = -(\hat{m}_o(X_t, Y_t^o, \sigma))'\Sigma^{-1}J + C.$$

Define $\pi_1 = \frac{\partial\pi_o}{\partial(Y_t^o)'}$ and take a differential:

$$\partial\pi_1 = -\partial\varepsilon_t^1\Sigma^{-1}J - \varepsilon_t^{'}\Sigma^{-1}\partial J + \partial C$$

As before, we have $\partial\varepsilon_t = J\partial Y_t^o$. So, we can write:

$$\partial\pi_1 = -(\partial Y_t^o)^{'} J^{'}\Sigma^{-1}J - \varepsilon_t^{'}\Sigma^{-1}\partial J + \partial C \tag{6.1}$$

From $\partial J = \sum_{j=1}^{n_o} F_j \partial Y_{t,j}^o$, we can write:

$$-\varepsilon_t^{'}\Sigma^{-1}\partial J = -\sum_{j=1}^{n_o} \varepsilon_t^{'}\Sigma^{-1}F_j\partial Y_{t,j}^o = -(\partial Y_t^o)^{'} \begin{pmatrix} \varepsilon_t^{'}\Sigma^{-1}F_1 \\ \varepsilon_t^{'}\Sigma^{-1}F_2 \\ \vdots \\ \varepsilon_t^{'}\Sigma^{-1}F_{n_o} \end{pmatrix} = (\partial Y_t^o)^{'} \begin{pmatrix} V_1^{'} \\ V_2^{'} \\ \vdots \\ V_{n_o}^{'} \end{pmatrix} = (\partial Y_t^o)^{'} V^{'}$$

Therefore, from (6.1) we can write $\partial\pi_1$ as:

$$\partial\pi_1 = (\partial Y_t^o)^{'}(-J^{'}\Sigma^{-1}J + V) + \partial C \tag{6.2}$$

where we have used the fact that V is symmetric. In order to calculate ∂C, first note that:

$$\partial(tr(J^{-1}F_i)) = tr(\partial J^{-1}F_i + J^{-1}\partial F_i)$$

and also that (e.g. Magnus and Neudecker (1999)):

$$\partial J^{-1} = -J^{-1}\partial J J^{-1}$$

such that, using $\partial J = \sum_{j=1}^{n_o} \frac{F_j}{\partial Y_{t,j}^o}$, we have that:

$$\partial(tr(J^{-1}F_i)) = \sum_{j=1}^{n_o} tr(-J^{-1}F_jJ^{-1}F_i)\partial Y_{t,j}^o + tr(J^{-1}\partial F_i)$$

Because $C = (tr(J^{-1}F_1)\dots tr(J^{-1}F_{n_o}))$ we have that:

$$\partial C = -(\partial Y_t^o)^{'}\partial F_i$$

Now using that $\partial F_i = \sum_{j=1}^{n_o} F_{ij}\partial Y_{t,j}^o$ we can write:

$$\partial C = -(\partial Y_t^o)^{'}A + (\partial Y_t^o)^{'}F \tag{6.3}$$

Combining (6.2) and (6.3) we get:

$$\partial\pi_1 = (\partial Y_t^o)^{'}(-J^{'}\Sigma^{-1}J + V - A + F)$$

which implies that the Hessian in $(-J'\Sigma^{-1}J + V - A + F)$, as we wanted to prove. $\square$

Proof of Proposition 3.2

Proof. To find the Jacobian let us take the differential of (3.13) which is given by:

$$\partial\varepsilon_t = \tilde{G}_o^{1,0}\partial Y_t^o + \tilde{G}_o^{2,0}(\partial Y_t^o \otimes Y_t^o) + \tilde{G}_o^{2,0}(Y_t^o \otimes \partial Y_t^o) + \tilde{G}_o^{1,1}(\partial Y_t^o \otimes X_t) \tag{6.4}$$

Note that $\tilde{G}_o^{2,0}$ contains second-order derivatives. In particular, each row of $\tilde{G}_o^{2,0}$ is the vectorized version of a Hessian matrix, which is symmetric. From this we have that:

$$\tilde{G}_o^{2,0}(\partial Y_t^o \otimes Y_t^o) = \tilde{G}_o^{2,0}(Y_t^o \otimes \partial Y_t^o),$$

so that (6.4) can be written as:

$$\begin{aligned}\partial\varepsilon_t &= \tilde{G}_o^{1,0}\partial Y_t^o + 2\tilde{G}_o^{2,0}(\partial Y_t^o \otimes Y_t^o) + \tilde{G}_o^{1,1}(\partial Y_t^o \otimes X_t) = \\ &= \tilde{G}_o^{1,0}\partial Y_t^o + 2\tilde{G}_o^{2,0}(I_{n_o} \otimes Y_t^o)\partial Y_t^o + \tilde{G}_o^{1,1}(I_{n_o} \otimes X_t)\partial Y_t^o \\ &= \big(\tilde{G}_o^{1,0} + 2\tilde{G}_o^{2,0}(I_{n_o} \otimes Y_t^o) + \tilde{G}_o^{1,1}(I_{n_o} \otimes X_t)\big)\partial Y_t^o\end{aligned} \tag{6.5}$$

This shows that the Jacobian is the expression given in (3.14):

$$J = \tilde{G}_o^{1,0} + 2\tilde{G}^{2,0}(I_{n_o} \otimes Y_t^o) + \tilde{G}_o^{1,1}(I_{n_o} \otimes X_t)$$

To calculate $(F_1, \ldots, F_{n_o})$ let us obtain the differential of J as:

$$\partial J = 2\tilde{G}^{2,0}(I_{n_o} \otimes \partial Y_t^o),$$

which shows that $F_j = 2\tilde{G}^{2,0}(I_{n_o} \otimes i_j)$ for $j = 1, \ldots, n_o$. Using (3.8) and (3.9) in Proposition 3.1 we obtain (3.15) and (3.16). The expressions for A and V in (3.17) and (3.18) were obtained by using the expression for F_j in equation (3.10) and (3.11) of Proposition 3.1. Finally, (3.19) is obtained from (3.12) by noting that F_j does not depend on Y_t^o, and, thus, $F_{ij} = 0$.

$\square$

Proof of Proposition 3.3

Proof. Here we use $\varepsilon_t = m_o(X_t, Y_t^o, \sigma)$ as the local inverse of the policy function $Y_t^o = g_o(X_t, \varepsilon_t, \sigma)$, which exists provided that the Jacobian is different from 0. Similarly, $Y_t^n = m_n(X_t, Y_t^o, \sigma)$ is the local inverse of $Y_t^n = g_n(X_t, \varepsilon_t, \sigma)$. From the properties of the inverse function, we have that $\varepsilon_t = m_o(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$, and also that $Y_t^n = m_n(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma) = g_n(X_t, \varepsilon_t, \sigma)$. Therefore, a second-order Taylor expansion of the composition functions $m_o(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$, and $m_n(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$ gives the following:

$$\begin{aligned} m_o(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma) = F_o^{0,0} + F_o^{1,0}\varepsilon_t + F_o^{0,1}X_t + F_o^{2,0}(\varepsilon_t \otimes \varepsilon_t) + \\ F_o^{0,2}(X_t \otimes X_t) + F_o^{1,1}(\varepsilon_t \otimes X_t) = \varepsilon_t \end{aligned} \tag{6.6}$$

$$\begin{aligned} m_n(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma) = F_n^{0,0} + F_n^{1,0}\varepsilon_t + F_n^{0,1}X_t + F_n^{2,0}(\varepsilon_t \otimes \varepsilon_t) + \\ F_n^{0,2}(X_t \otimes X_t) + F_n^{1,1}(\varepsilon_t \otimes X_t) = Y_n = g_n(X_t, \varepsilon_t, \sigma) \end{aligned} \tag{6.7}$$

Equations (6.6)-(6.7) imply the following restrictions on the F matrices:

$$\begin{aligned} &F_o^{0,0} = 0, \quad F_o^{1,0} = I, \quad F_o^{0,1} = 0, \quad F_o^{2,0} = 0, \quad F_o^{0,2} = 0, \quad F_o^{1,1} = 0, \\ &F_n^{0,0} = G_n^{0,0}, \quad F_n^{1,0} = G_n^{1,0}, \quad F_n^{0,1} = G_n^{0,1}, \quad F_n^{2,0} = G_n^{2,0}, \\ &F_n^{0,2} = G_n^{0,2}, \quad F_n^{1,1} = G_n^{1,1} \end{aligned} \tag{6.8}$$

Let $\tilde{g}_{o,s}(X_t, \varepsilon_t, \sigma)$ and $\tilde{g}_{n,s}(X_t, \varepsilon_t, \sigma)$ be the Taylor approximations of order s to the policy functions $g_o(X_t, \varepsilon_t, \sigma)$ and $g_n(X_t, \varepsilon_t, \sigma)$, respectively. The second-order Taylor approximation of the composition functions $m_o(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$, and $m_n(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$ can be obtained by calculating the composition function of the corresponding Taylor polynomials and keeping the terms up to 2^{nd} order.

Therefore, using the Taylor approximations in (2.4) and in (3.20) we can calculate the compositions $m_o(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$, and $m_n(X_t, g_o(X_t, \varepsilon_t, \sigma), \sigma)$, and obtain the coefficients for $(\varepsilon_t \otimes \varepsilon_t)$, (ε_t), $(\varepsilon_t \otimes X_t)$, $(X_t \otimes X_t)$, (X_t) and for the constant term. Equating these coefficients with those in (6.8) gives the following 12 equations with 12 unknowns.

for $(\varepsilon_t \otimes \varepsilon_t)$

$$\begin{cases} \tilde{G_o}^{2,0}(G_o^{1,0} \otimes G_o^{1,0}) + 2\tilde{G_o}^{2,0}(G_o^{0,0} \otimes G_o^{2,0}) + \tilde{G_o}^{1,0}G_o^{2,0} = F_o^{2,0} = 0 \\ \\ \tilde{G_n}^{2,0}(G_n^{1,0} \otimes G_n^{1,0}) + 2\tilde{G_n}^{2,0}(G_n^{0,0} \otimes G_n^{2,0}) + \tilde{G_n}^{1,0}G_n^{2,0} = F_n^{2,0} = G_n^{2,0} \end{cases} \tag{6.9}$$

for (ε_t)

$$\begin{cases} \tilde{G_o}^{1,0}G_o^{1,0} + 2\tilde{G_o}^{2,0}(G_o^{0,0} \otimes G_o^{1,0}) = F_o^{1,0} = I \\ \\ \tilde{G_n}^{1,0}G_o^{1,0} + 2\tilde{G_n}^{2,0}(G_o^{0,0} \otimes G_o^{1,0}) = F_n^{1,0} = G_n^{1,0} \end{cases} \tag{6.10}$$

for $(\varepsilon_t \otimes X_t)$

$$\begin{cases} \tilde{G_o}^{1,1}(G_o^{1,0} \otimes I_K) + \tilde{G_o}^{1,0} G_o^{1,1} + 2\tilde{G_o}^{2,0}(G_o^{1,0} \otimes G_o^{0,1}) + 2\tilde{G_o}^{2,0}(G_o^{1,1} \otimes G_o^{0,0}) = F_o^{1,1} = 0 \\ \tilde{G_n}^{1,1}(G_o^{1,0} \otimes I_K) + \tilde{G_n}^{1,0} G_o^{1,1} + 2\tilde{G_n}^{2,0}(G_o^{1,0} \otimes G_o^{0,1}) + 2\tilde{G_n}^{2,0}(G_o^{1,1} \otimes G_o^{0,0}) = F_n^{1,1} = G_n^{1,1} \end{cases} \tag{6.11}$$

for $(X_t \otimes X_t)$

$$\begin{cases} \tilde{G_o}^{0,2} + \tilde{G_o}^{1,0} G_o^{0,2} + \tilde{G_o}^{2,0}(G_o^{0,1} \otimes G_o^{0,1}) + \tilde{G_o}^{1,1}(G_o^{1,0} \otimes I_K) + 2\tilde{G_o}^{2,0}(G_o^{0,0} \otimes G_o^{0,2}) = F_o^{0,2} = 0 \\ \tilde{G_n}^{0,2} + \tilde{G_n}^{1,0} G_o^{0,2} + \tilde{G_n}^{2,0}(G_o^{0,1} \otimes G_o^{0,1}) + \tilde{G_n}^{1,1}(G_o^{1,0} \otimes I_K) + 2\tilde{G_n}^{2,0}(G_o^{0,0} \otimes G_o^{0,2}) = F_n^{0,2} = G_n^{0,2} \end{cases} \tag{6.12}$$

for (X_t)

$$\begin{cases} \tilde{G_o}^{0,1} + \tilde{G_o}^{1,0} G_o^{0,1} + \tilde{G_o}^{1,1}(G_o^{0,0} \otimes I_K) + 2\tilde{G_o}^{2,0}(G_o^{0,0} \otimes G_o^{0,1}) = F_o^{0,1} = 0 \\ \tilde{G_n}^{0,1} + \tilde{G_n}^{1,0} G_o^{0,1} + \tilde{G_n}^{1,1}(G_o^{0,0} \otimes I_K) + 2\tilde{G_n}^{2,0}(G_o^{0,0} \otimes G_o^{0,1}) = F_n^{0,1} = G_n^{0,1} \end{cases} \tag{6.13}$$

for the constant term:

$$\begin{cases} \tilde{G_o}^{0,0} + \tilde{G_o}^{1,0} G_o^{0,0} + \tilde{G_o}^{2,0}(G_o^{0,0} \otimes G_o^{0,0}) = F_o^{0,0} = 0 \\ \tilde{G_n}^{0,0} + \tilde{G_n}^{1,0} G_o^{0,0} + \tilde{G_n}^{2,0}(G_o^{0,0} \otimes G_o^{0,0}) = F_n^{0,0} = G_n^{0,0} \end{cases} \tag{6.14}$$

Thus, solving the system of equations in (6.9)-(6.14) through substituting and collecting terms we obtain the proposed solutions for matrices of the second-order approximation of the inverses. □

Appendix B: Additional Tables and Figures for the Empirical Analysis.

Name	**Model equation**

FOC consumption:	$$d_t\left(\tilde{c}_t - h\tilde{c}_{t-1}\frac{z_{t-1}}{z_t}\right)^{-1} - h\beta E_t d_{t+1}\left(\tilde{c}_{t+1}\frac{z_{t+1}}{z_t} - h\tilde{c}_t\right)^{-1} = \tilde{\lambda}_t \quad (6.1)$$
Euler equation:	$$\tilde{\lambda}_t = \beta E_t\{\tilde{\lambda}_{t+1}\frac{z_t}{z_{t+1}}\frac{R_t}{\Pi_{t+1}}\} \quad (6.2)$$
FOC capital utilization:	$$\tilde{r}_t = \gamma_1 + \gamma_2(u_t - 1) \quad (6.3)$$
FOC capital:	$$\begin{aligned}\tilde{q}_t = \beta E_t\Big\{&\frac{\tilde{\lambda}_{t+1}}{\tilde{\lambda}_t}\frac{z_t}{z_{t+1}}\frac{\mu_t}{\mu_{t+1}}((1-\delta)\tilde{q}_{t+1} + \tilde{r}_{t+1}u_{t+1} - \\ &- (\gamma_1(u_{t+1}-1) + \frac{\gamma_2}{2}(u_{t+1}-1)^2))\Big\}\end{aligned} \quad (6.4)$$
FOC investment:	$$\begin{aligned}1 = \tilde{q}_t&\left(1 - S\left[\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right] - S'\left[\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right]\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\frac{\tilde{z}_t}{\tilde{z}_{t-1}}\right) + \\ &+\beta E_t\tilde{q}_{t+1}\frac{\tilde{\lambda}_{t+1}}{\tilde{\lambda}_t}\frac{z_t}{z_{t+1}}S'\left[\frac{\tilde{x}_{t+1}}{\tilde{x}_t}\frac{z_{t+1}}{z_t}\right]\left(\frac{\tilde{x}_{t+1}}{\tilde{x}_t}\frac{\tilde{z}_{t+1}}{\tilde{z}_t}\right)^2\end{aligned} \quad (6.5)$$
	where: $$S\left[\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right] = \frac{\kappa}{2}\left(\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\frac{z_t}{z_{t-1}} - \exp\Lambda_z\right)^2$$ $$S'\left[\frac{\tilde{x}_{t+1}}{\tilde{x}_t}\frac{z_{t+1}}{z_t}\right] = \kappa\left(\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\frac{z_t}{z_{t-1}} - \exp\Lambda_z\right)$$

Wage setting 1:	$$f_t = \frac{\eta-1}{\eta}(\tilde{\omega}_t^*)^{1-\eta}\tilde{\lambda}_t(\tilde{\omega}_t)^{\eta}l_t^d + + \beta\theta_\omega E_t\left(\frac{\Pi_t^{\chi_\omega}}{\Pi_{t+1}}\right)^{1-\eta}\left(\frac{\tilde{\omega}_{t+1}^*}{\tilde{\omega}_t^*}\frac{z_{t+1}}{z_t}\right)^{\eta-1}f_{t+1} \quad (6.6)$$
Wage setting 2:	$$f_t = \psi d_t\varphi_t(\Pi_t^{*\omega})^{-\eta(1+\gamma)}(l_t^d)^{1+\gamma} + + \beta\theta_\omega E_t\left(\frac{\Pi_t^{\chi_\omega}}{\Pi_{t+1}}\right)^{\eta(1+\gamma)}\left(\frac{\tilde{\omega}_{t+1}^*}{\tilde{\omega}_t^*}\frac{z_{t+1}}{z_t}\right)^{\eta(1+\gamma)}f_{t+1} \quad (6.7)$$
Firm price setting 1:	$$g_t^1 = \tilde{\lambda}_t mc_t\tilde{y}_t^d + \beta\theta_p E_t\left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{-\varepsilon}g_{t+1}^1 \quad (6.8)$$
Firm price setting 2:	$$g_t^2 = \tilde{\lambda}_t\Pi_t^*\tilde{y}_t^d + \beta\theta_p E_t\left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{1-\varepsilon}\left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right)g_{t+1}^2 \quad (6.9)$$
Firm price setting 3:	$$\varepsilon g_t^1 = (\varepsilon-1)g_t^2 \quad (6.10)$$
Optimal capital labor ratio:	$$\frac{u_t\tilde{k}_{t-1}}{l_t^d} = \frac{\alpha}{(1-\alpha)}\frac{\tilde{\omega}_t}{\tilde{r}_t}\frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}} \quad (6.11)$$
Marginal costs:	$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}(\tilde{\omega}_t)^{1-\alpha}\tilde{r}_t^{\alpha} \quad (6.12)$$

Law of motion wages:	$$1 = \theta_\omega \left(\frac{\Pi_{t-1}^{\chi_\omega}}{\Pi_t}\right)^{1-\eta} \left(\frac{\tilde{\omega}_{t-1}}{\tilde{\omega}_t}\frac{z_{t-1}}{z_t}\right)^{1-\eta} + (1-\theta_\omega)\left(\Pi_t^{*\omega}\right)^{1-\eta} \quad (6.13)$$
Law of motion prices:	$$1 = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon} + (1-\theta_p)\Pi_t^{*1-\varepsilon} \quad (6.14)$$
Taylor Rule:	$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{\frac{\tilde{y}_t^d}{\tilde{y}_{t-1}^d}\frac{z_t}{z_{t-1}}}{\Lambda_{y^d}}\right)^{\gamma_y}\right)^{1-\gamma_R} \exp\left(m_t\right) \quad (6.15)$$
Resource constraint:	$$\tilde{y}_t^d = \tilde{c}_t + \tilde{x}_t + \frac{z_{t-1}}{z_t}\frac{\mu_{t-1}}{\mu_t}\left(\gamma_1(u_t-1) + \frac{\gamma_2}{2}(u_t-1)^2\right)\tilde{k}_{t-1} \quad (6.16)$$
Aggregate production:	$$\tilde{y}_t^d = \frac{\frac{A_t}{A_{t-1}}\frac{z_{t-1}}{z_t}\left(u_t\tilde{k}_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha} - \phi}{v_t^p} \quad (6.17)$$
Aggregate labor market:	$$l_t = v_t^\omega l_t^d \quad (6.18)$$
LOM Price dispersion term:	$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + (1-\theta_p)\Pi_t^{*-\varepsilon} \quad (6.19)$$

LOM Wage dispersion term:	$$v_t^w = \theta_w \left(\frac{\tilde{\omega}_{t-1}}{\tilde{\omega}_t} \frac{\Pi_{t-1}^{\chi_w}}{\Pi_t} \right)^{-\eta} v_{t-1}^w + (1-\theta_w)(\Pi_t^{*w})^{-\eta} \quad (6.20)$$
Law of motion for capital:	$$\begin{aligned} &\tilde{k}_t \frac{z_t}{z_{t-1}} \frac{\mu_t}{\mu_{t-1}} - (1-\delta)\tilde{k}_{t-1} - \\ &\frac{z_t}{z_{t-1}} \frac{\mu_t}{\mu_{t-1}} \left(1 - S\left[\frac{\tilde{x}_t}{\tilde{x}_{t-1}} \frac{z_t}{z_{t-1}}\right]\right) \tilde{x}_t = 0 \end{aligned} \quad (6.21)$$
Profits:	$$F_t = \tilde{y}_t^d - \frac{1}{(1-\alpha)} \tilde{\omega}_t l_t^d \quad (6.22)$$
Definition optimal wage inflation:	$$\Pi_t^{*\omega} = \frac{\omega_t^*}{\omega_t}, \text{ where } \omega_t^* = \tilde{\omega}_t^* z_t, \ \omega_t = \tilde{\omega}_t z_t \quad (6.23)$$
Preference Shock:	$$\log d_t = \rho^d \log d_{t-1} + \varepsilon_{d,t} \quad (6.24)$$
Labor disutility Shock:	$$\log \varphi_t = \rho^\varphi \log \varphi_{t-1} + \varepsilon_{\varphi,t} \quad (6.25)$$
Investment specific technology:	$$\log \mu_{I,t} = \Lambda_\mu + \varepsilon_{\mu_I,t}, \text{ where } \mu_{I,t} = \frac{\mu_t}{\mu_{t-1}} \quad (6.26)$$
Neutral technology:	$$\log \mu_{A,t} = \Lambda_A + \varepsilon_{A,t}, \text{ where } \mu_{A,t} = \frac{A_t}{A_{t-1}} \quad (6.27)$$

Definition composite technology:	$$\mu_{z,t} = \mu_{A,t}^{\frac{1}{(1-\alpha)}} \mu_{I,t}^{\frac{\alpha}{(1-\alpha)}}, \text{ where } \mu_{z,t} = \frac{z_t}{z_{t-1}} \tag{6.28}$$
and	$$\gamma_1 = \mu_{z,t} \frac{\mu_{I,t}}{\beta} - (1-\delta) \tag{6.29}$$
and	$$\Lambda_x = \exp \Lambda_z \tag{6.30}$$
and	$$\Lambda_z = \frac{\Lambda_A + \alpha \Lambda_\mu}{1-\alpha} \tag{6.31}$$
observation equation 1	$$y_t^o = \log(\tilde{y}_t^d) - \log(\tilde{y}_{t-1}^d) + \log(\mu_{z,t}) \tag{6.32}$$
observation equation 2	$$\omega_t^o = \log(\tilde{\omega}_t) - \log(\tilde{\omega}_{t-1}) + \log(\mu_{z,t}) \tag{6.33}$$

Table B1: Model Equilibrium Equations

Variable	**Description**	**Variable**	**Description**
$\tilde{c}_t$	consumption	Π	Inflation
d_t	shock to intertemporal preferences	$\tilde{\lambda}_t$	Lagrange multiplier
$\mu_{z,t}$	trend growth rate of the economy	$\mu_{I,t}$	growth rate of investment-specific technology growth
$\mu_{A,t}$	growth rate of neutral technology	R_t	Nominal Interest rate
$\tilde{r}_t$	rental rate of capital	$\tilde{x}_t$	investment
u_t	capacity utilization	$\tilde{q}_t$	Tobin marginal q
f_t	recursive formulation of wage setting	l_t^d	aggregate labor demand
$\tilde{\omega}_t$	real wage	$\tilde{\omega}_t^*$	optimal real wage
Π_t^*	optimal price inflation	$\Pi_t^{\omega*}$	optimal wage inflation
$\tilde{y}_t^d$	aggregate output	mc_t	marginal costs
k_t	capital	l_t	aggregate labor bundle
g_t^1	variable 1 for recursive formulation of price setting	g_t^2	variable 2 for recursive formulation of price setting
v_t^p	price dispersion term	v_t^ω	wage dispersion term
φ_t	labor disutility shock	F_t	firm profits
ω_t	non-detrended real wage	ω_t^*	non-detrended optimal real wage

Table B2: Variables in the Model.

Name in Database	Transformation	Model Notation	Description
Relative Price of Investment Goods (PIRIC)	$-\Delta \log(x)$	$\log(\mu_{I,t})$	log of growth rate of investment specific technology growth
Real gross domestic product per capita (A939RX0Q 048SBEA)	$\Delta \log(x)$	y_t^o	real output per capita growth
Nonfarm Business Sector: Real Compensation Per Hour (COMPRNFB)	$\Delta \log(x)$	ω_t^o	real wages per capita growth
Gross Domestic Product: Implicit Price Deflator (GDPDEF)	$\Delta \log(x)$	Π_t	log of gross inflation
Effective Federal Funds Rate (FEDFUNDS)	$\log(1 + x/400)$	R_t	log of gross nominal interest rate

Table B3: Data Sources and Transformations Used. All variables were obtained from the Federal Reserve Bank of St. Louis' FRED database. The column 'Transformation' indicates the transformation to the original series in the database to match the model variable indicated in the column 'Model Notation'. y_t^o and ω_t^o are defined in the observation equations (6.32)-(6.33) in Table B1.

	Distr.	Prior distribution		Posterior distribution			
				M_1 - **First-order**		M_2 - **Second-order**	
		Mean	**St.Dev.**	**Mean**	**Credible interval**	**Mean**	**Credible interval**
β	Gamma	0.998	0.1	0.9983	[0.9972, 0.9992]	0.9990	[0.9983, 0.9996]
h	Beta	0.7	0.1	0.5523	[0.4448, 0.6488]	0.8382	[0.8149, 0.8592]
ψ	Normal	9	3	9.0389	[3.4819, 14.8717]	9.8073	[5.7470, 15.1635]
γ	Normal	1	0.25	0.0761	[-0.0775, 0.3174]	1.6953	[1.2575, 2.1099]
κ	Normal	4	1.5	5.8466	[3.6256, 8.3106]	0.2005	[0.1432, 0.2709]
α	Normal	0.3	0.025	0.2937	[0.2521, 0.3351]	0.3069	[0.2643, 0.3501]
θ_p	Beta	0.5	0.1	0.6466	[0.5714, 0.7114]	0.5948	[0.5033, 0.6539]
χ	Beta	0.5	0.1	0.1228	[0.0433, 0.2403]	0.3335	[0.1624, 0.5760]
θ_ω	Beta	0.5	0.1	0.3361	[0.2166, 0.5528]	0.3138	[0.2693, 0.3598]
χ_ω	Beta	0.5	0.1	0.5408	[0.3551, 0.7189]	0.4496	[0.2732, 0.6178]
γ_R	Beta	0.75	0.1	0.6830	[0.6276, 0.7297]	0.7086	[0.6600, 0.7516]
γ_Y	Normal	0.120	0.05	0.1791	[0.0962, 0.2662]	0.2971	[0.2376, 0.3557]
γ_π	Normal	1.5	0.125	1.6409	[1.4678, 1.8212]	1.7790	[1.6430, 1.9509]
$\bar{\Pi}$	Gamma	1.01	0.1	1.0089	[1.0075, 1.0103]	1.0082	[1.0071, 1.0095]
ρ_d	Beta	0.5	0.2	0.9140	[0.8672, 0.9572]	0.7451	[0.6844, 0.8066]
ρ_ϕ	Beta	0.5	0.2	0.9948	[0.9892, 0.9989]	0.9931	[0.9860, 0.9984]

Table B4: Prior and Posterior Distribution for Structural Parameters of M_1 and M_2 (Part I)

	Distr.	Prior distribution		Posterior distribution			
				M_1 - **First-order**		M_2 - **Second-order**	
		Mean	**St.Dev.**	**Mean**	**Credible interval**	**Mean**	**Credible interval**
Λ_μ	Normal	0.34	0.1	0.0056	[0.0049, 0.0063]	0.0059	[0.0051, 0.0065]
Λ_A	Normal	0.178	0.075	0.0007	[-0.0002, 0.0016]	0.0010	[0.0001, 0.0018]
γ_2	Beta	0.01	0.03	0.2704	[0.1168, 0.4904]	0.2768	[0.1075, 0.4830]
δ	Beta	0.025	0.015	0.0652	[0.0344, 0.1108]	0.0396	[0.0331, 0.0517]
σ_A	IG*	0.1	2	-4.4390	[-4.5647, -4.3140]	-4.5792	[-4.7789, -4.3977]
σ_d	IG*	0.1	2	-3.6454	[-3.9580, -3.2126]	-2.5250	[-2.7005, -2.3576]
σ_φ	IG*	0.1	2	-3.6866	[-3.9529, -3.4014]	-2.4428	[-2.6186, -2.2635]
σ_μ	IG*	0.1	2	-5.0890	[-5.1792, -4.9960]	-5.0900	[-5.1741, -5.0024]
σ_m	IG*	0.1	2	-5.7512	[-5.8581, -5.6345]	-5.7818	[-5.8811, -5.6742]

Table B5: Prior and Posterior Distribution for Structural Parameters of M_1 and M_2 (Part II). *The prior distribution is given for $\exp \sigma_i$ and the posterior for σ_i. IG denotes the Inverse Gamma distribution.

	Distr.	Prior distribution		Posterior distribution			
		Mean	St.Dev.	$M_{2,G}$ - **GARCH**		$M_{2,fG}$ - **fGARCH**	
				Mean	**Credible interval**	**Mean**	**Credible interval**
β	Gamma	0.998	0.1	0.9987	[0.9977, 0.9994]	0.9986	[0.9977, 0.9984]
h	Beta	0.7	0.1	0.5525	[0.4622, 0.6492]	0.5645	[0.4695, 0.6619]
ψ	Normal	9	3	9.0195	[3.4322, 14.8874]	9.0756	[3.3670, 14.8716]
γ	Normal	1	0.25	-0.0382	[-0.0653, -0.0061]	0.0665	[-0.0866, 0.3564]
κ	Normal	4	1.5	5.5269	[3.0775, 8.2445]	6.2402	[3.9788, 8.7965]
α	Normal	0.3	0.025	0.2730	[0.2322, 0.3119]	0.2846	[0.2425, 0.3257]
θ_p	Beta	0.5	0.1	0.5598	[0.5059, 0.6004]	0.5577	[0.4911, 0.6144]
χ	Beta	0.5	0.1	0.2181	[0.1669, 0.2582]	0.1574	[0.0584, 0.2929]
θ_ω	Beta	0.5	0.1	0.3353	[0.2936, 0.3809]	0.2898	[0.1424, 0.6048]
χ_ω	Beta	0.5	0.1	0.5171	[0.3225, 0.7110]	0.4805	[0.2884, 0.6749]
γ_R	Beta	0.75	0.1	0.6872	[0.6468, 0.7187]	0.8038	[0.7687, 0.8349]
γ_Y	Normal	0.120	0.05	0.1688	[0.1094, 0.2217]	0.1770	[0.0937, 0.2617]
γ_π	Normal	1.5	0.125	1.6800	[1.5929, 1.7591]	1.6974	[1.5432, 1.8576]
$\bar{\Pi}$	Gamma	1.01	0.1	1.0073	[1.0063, 1.0083]	1.0064	[1.0057, 1.0071]
ρ_d	Beta	0.5	0.2	0.9313	[0.9172, 0.9401]	0.9152	[0.8830, 0.9436]
ρ_ϕ	Beta	0.5	0.2	0.9954	[0.9912, 0.9989]	0.9954	[0.9911, 0.9987]

Table B6: Prior and Posterior Distribution for Structural Parameters of $M_{2,G}$ and $M_{2,fG}$ (Part I).

	Distr.	Prior distribution		Posterior distribution			
				$M_{2,G}$ - **GARCH**		$M_{2,fG}$ - **fGARCH**	
		Mean	**St.Dev.**	**Mean**	**Credible interval**	**Mean**	**Credible interval**
Λ_μ	Normal	0.34	0.1	0.0054	[0.0048, 0.0060]	0.0058	[0.0051, 0.0064]
Λ_A	Normal	0.178	0.075	0.0006	[-0.0002, 0.0014]	0.0003	[-0.0005, 0.0010]
ρ_1^d	Beta	0.7	0.046	0.6533	[0.6101, 0.6886]	0.6019	[0.4979, 0.6550]
ρ_2^d	Beta	0.2	0.12	0.1377	[0.1243, 0.1552]	-	-
ρ_1^φ	Beta	0.7	0.046	0.6933	[0.5972, 0.7795]	-	-
ρ_2^φ	Beta	0.2	0.12	0.0568	[0.0159, 0.1048]	-	-
ρ_1^μ	Beta	0.7	0.046	0.6610	[0.5711, 0.8102]	-	-
ρ_2^μ	Beta	0.2	0.12	0.1377	[0.0843, 0.1956]	-	-
ρ_1^A	Beta	0.7	0.046	0.6974	[0.5934, 0.7877]	-	-
ρ_2^A	Beta	0.2	0.12	0.0810	[0.0251, 0.1448]	-	-
ρ_1^m	Beta	0.7	0.046	0.7194	[0.6216, 0.8128]	-	-
ρ_2^m	Beta	0.2	0.12	0.0483	[0.0321, 0.0680]	-	-
δ^d	Normal	0	0.2	-	-	-0.0392	[0.0274, 0.0517]
δ^φ	Normal	0	0.2	-	-	-0.0064	[-0.0021, 0.0160]
δ^μ	Normal	0	0.2	-	-	-0.1302	[0.1007, 0.1716]
δ^A	Normal	0	0.2	-	-	0.0217	[-0.0493, 0.0029]
δ^m	Normal	0	0.2	-	-	0.6096	[-0.6770, -0.5622]
γ_2	Beta	0.01	0.03	0.3002	[0.1550, 0.4980]	0.3381	[0.1654, 0.5544]
δ	Beta	0.025	0.015	0.1137	[0.0618, 0.1897]	0.0900	[0.0494, 0.1503]

Table B7: Prior and Posterior Distribution for Structural Parameters of $M_{2,G}$ and $M_{2,fG}$ (Part II)

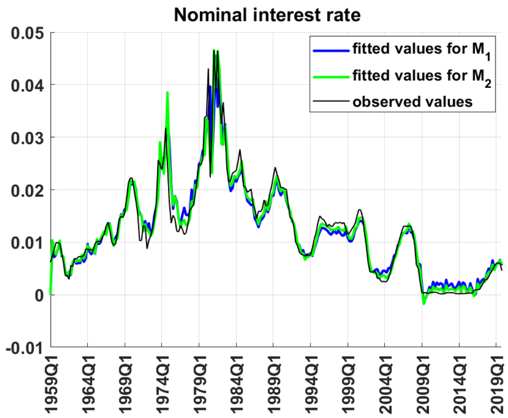
	Distr.	Prior distribution		Posterior distribution			
		Mean	St.Dev.	$M_{2,G}$ - **GARCH**		$M_{2,fG}$ - **fGARCH**	
				Mean	**Credible interval**	**Mean**	**Credible interval**
σ_A	IG*	0.1	2	-4.7088**	[-4.8680, -4.5454]	-2.9338	[-3.5957, -2.2367]
σ_d	IG*	0.1	2	-3.9732	[-4.0799, -3.8386]	-2.1175	[-2.7748, -1.4216]
σ_φ	IG*	0.1	2	-3.8355	[-4.000, -3.6370]	-1.9639	[-2.6691, -1.2032]
σ_μ	IG*	0.1	2	-5.0391	[-5.1912, -4.8591]	-3.3121	[-3.9670, -2.6109]
σ_m	IG*	0.1	2	-5.7833	[-5.9118, -5.6464]	-4.2818	[-4.9142, -3.5850]

Table B8: Prior and Posterior distribution for structural parameters of $M_{2,G}$ and $M_{2,fG}$ (Part III). *The prior distribution is given for $\exp \sigma_i$ and the posterior for σ_i. IG denotes the Inverse Gamma distribution.

Model variables	**Model** M_1	**Model** M_2	**Deviations**
Consumption, c	-2.068	-0.5128	-1.555
Investment, x	-3.123	-1.5038	-1.619
Labor demand, ld	-2.285	-0.9476	-1.338
Output, yd	-1.769	-0.1971	-1.572
Capital, k	-0.5121	1.5109	-2.022
Aggregate labor, l	-2.2851	-0.9473	-1.3378
Firm profits, F	-4.073	-2.4990	-1.5727

Table B9: Steady-State Values for Models M_1 and M_2.

42



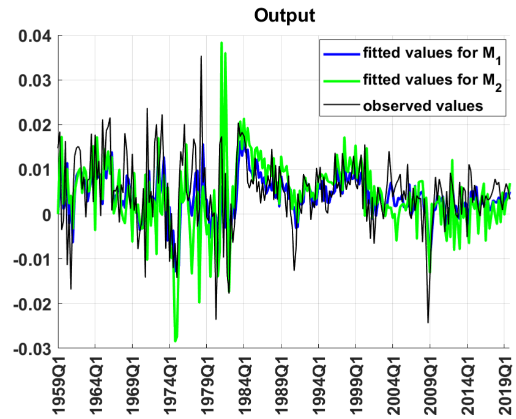


Figure B1: Fitted vs Observed Values for M_1 and M_2: Nominal Interest Rate and Output.

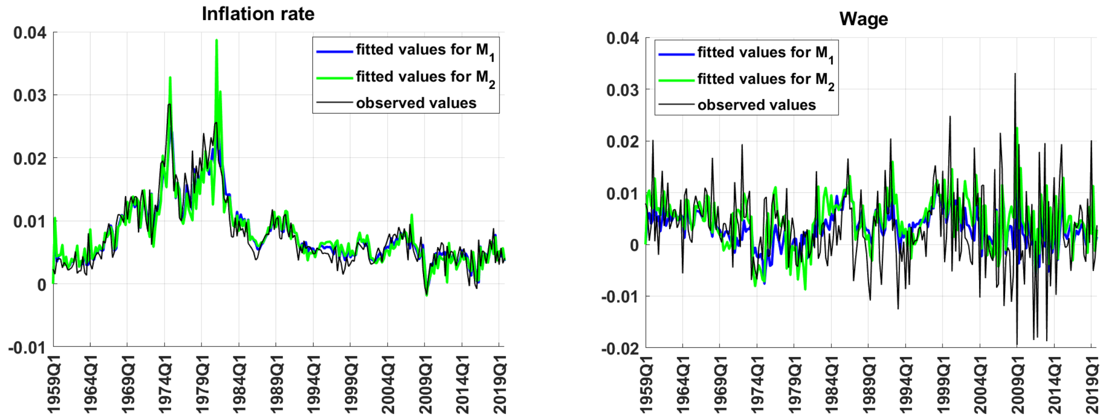


Figure B2: Fitted vs Observed Values for M_1 and M_2: Inflation and Wage.

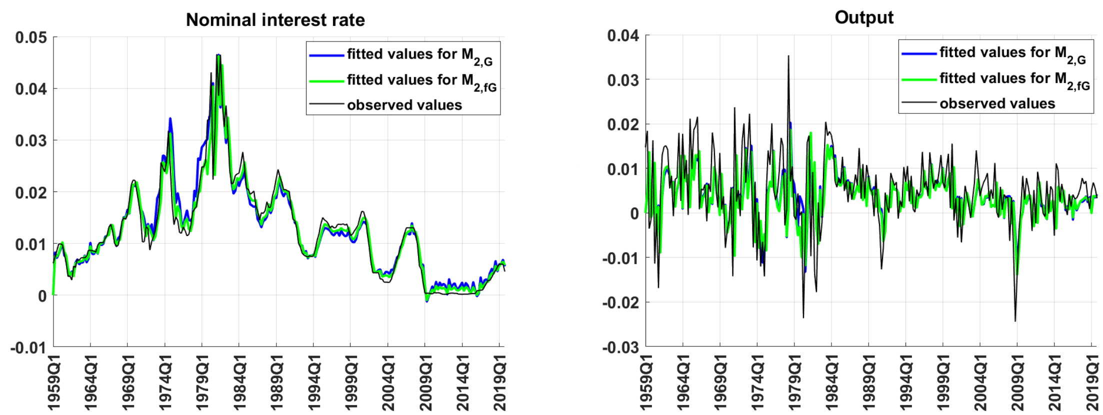


Figure B3: Fitted vs Observed Values for $M_{2,G}$ and $M_{2,fG}$: Nominal Interest Rate and Output.

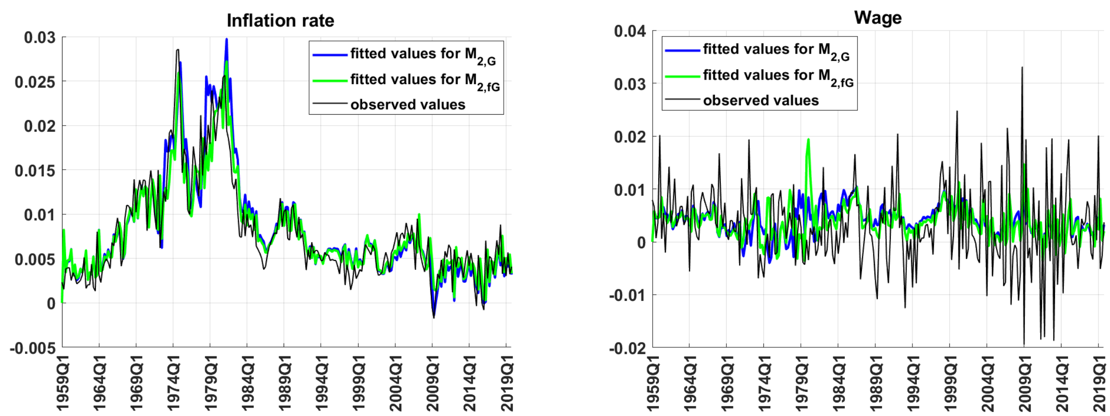


Figure B4: Fitted vs Observed values for $M_{2,G}$ and $M_{2,fG}$: Inflation and Wage.

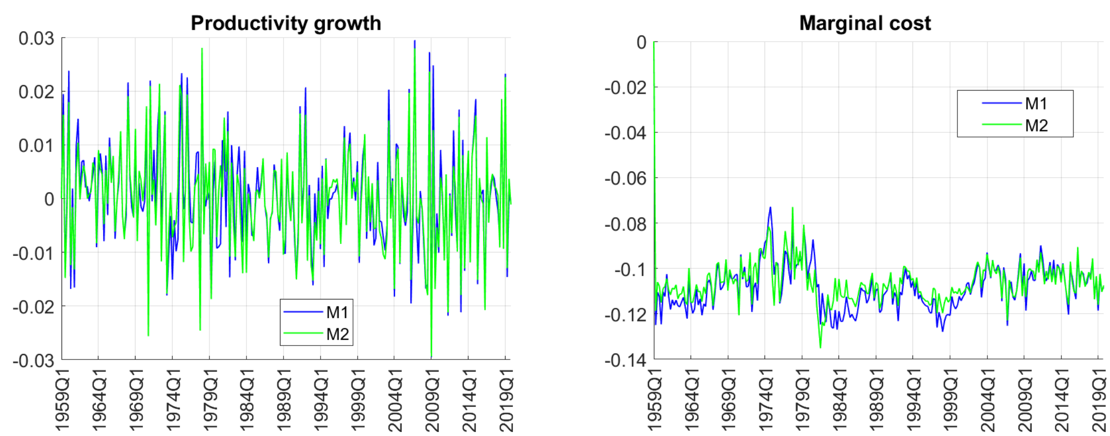


Figure B5: Posterior Estimates of Productivity Growth and Marginal Cost for M_1 and M_2.

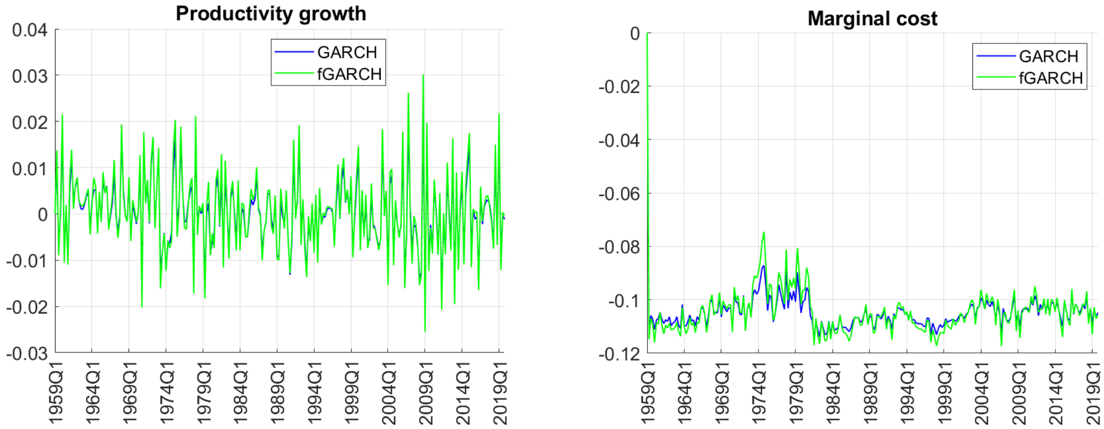


Figure B6: Posterior Estimates of Productivity Growth and Marginal Cost for $M_{2,G}$ and $M_{2,fG}$.

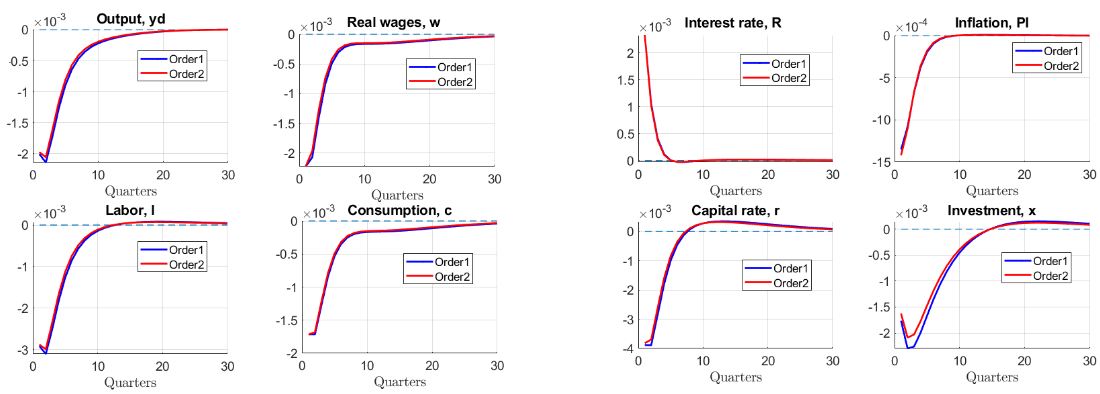


Figure B7: IRFs to a Monetary Policy Shocks for M_1 and M_2 Using the Same Values for the Parameters. The values of the parameters are the posterior estimates in the linear model.

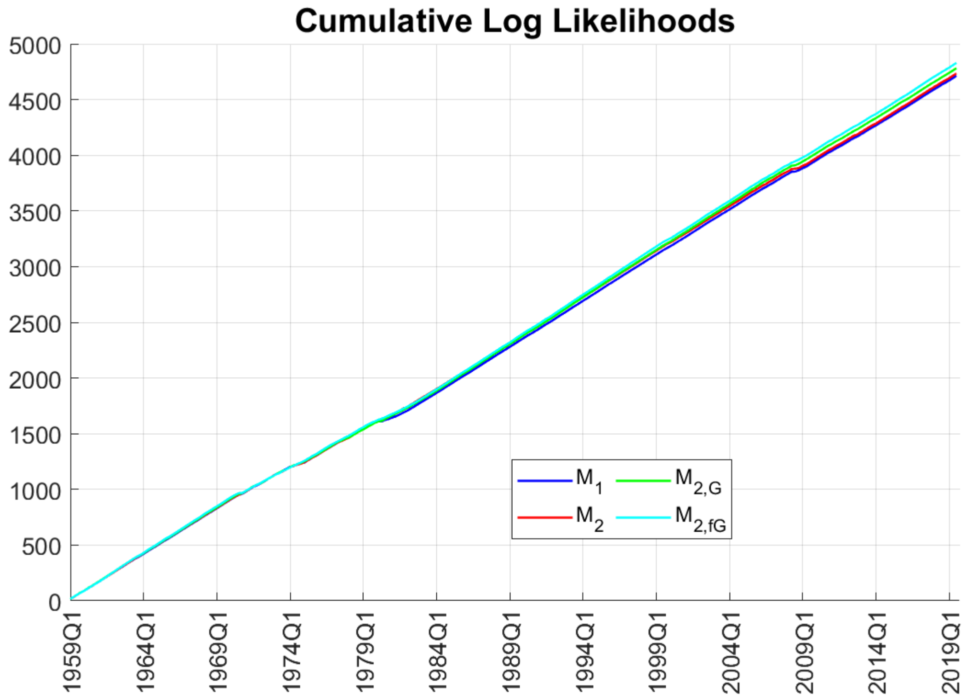


Figure B8: Cumulative Log Likelihoods for M_1, M_2, $M_{2,G}$ and $M_{2,fG}$.