



Documentos de trabajo

Economía y Finanzas

N° 20-07

2020

Confronting DSGE model with data

Alfredo Villca

Confrontando un modelo DSGE con los datos*

Alfredo Villca †

Diciembre, 2019

Abstract

En este documento intentamos responder a una cuestión clave en la modelación macroeconómica ¿Qué tanto reproduce un modelo estocástico de equilibrio general los datos? Para responder esta pregunta usamos tres tipos de técnicas. Primero, comparamos los momentos estadísticos del modelo con los datos. Segundo, efectuamos una simulación de los ciclos. Tercero, comparamos las impulso-respuesta del modelo DSGE con las de un modelo SVAR. Adicionalmente, se estima el modelo usando métodos bayesianos. Tomando en cuenta datos de Estados Unidos durante 1950-2019, el modelo replica relativamente bien el comportamiento de los datos, aunque se observan ciertas diferencias en cuanto a las magnitudes.

Clasificación JEL: E32, E37, C32, C11

Palabras Clave: Ciclos económicos, simulación de ciclos económicos, modelo SVAR, análisis bayesiano

*Este documento es producto de los cursos de macroeconomía avanzada del Doctorado en Economía de la Universidad EAFIT. Los conceptos expresados en este documento de trabajo son responsabilidad exclusiva del autor y no comprometen a la Universidad EAFIT ni al Centro de Investigaciones Económicas y Financieras (CIEF).

†Estudiante del programa de Doctorado en Economía, Universidad EAFIT. Grupo de Estudios en Economía y Empresa, Departamento de Economía, Centro de Investigaciones Económicas y Financieras (CIEF), Universidad EAFIT, Carrera 49 Número 7 Sur 50, Medellín, Colombia. Correo: avillca@eafit.edu.co

Confronting DSGE model with data

Abstract

In this document we try to answer a key question in macroeconomic modeling. How much does a stochastic model of general equilibrium reproduce the data? To answer this question we use three metrics. First, we compare the statistical moments of the model with the data. Second, we perform a simulation of the cycles. Third, we compare the impulse-response of the DSGE model with those of an SVAR model. Additionally, the model is estimated using Bayesian methods. Taking into account data from the United States, on a quarterly frequency (1950-2019), the model replicates the behavior of the data relatively well, although certain differences in magnitudes are observed.

Classification JEL: E32, E37, C32, C11

Key Words: Business cycles, Business cycle simulation, SVAR model, Bayesian analysis

1 Introducción

Los modelos dinámicos estocásticos de equilibrio general (DSGE, por sus siglas en inglés) se han convertido en una de las principales herramientas para determinar las fuentes de los ciclos económicos y la simulación de las implicancias de política económica. La base de estos modelos son los llamados modelos de Ciclos Económicos Reales (RBC, por sus siglas en inglés) desarrollados por [Kydland and Prescott \(1982\)](#), [Long and Plosser \(1983\)](#), [King et al. \(1988a\)](#), [King et al. \(1988b\)](#), [Plosser \(1989\)](#), entre otros. Estos modelos han transitado por muchos cambios con distintas extensiones. Desde aquellos que incorporan otros agentes hasta aquellos que han incorporado diferentes fricciones y rigideces en los mercados.

Sin embargo, en la mayoría de los casos se analizan las funciones de impulso respuesta. En pocas ocasiones se han confrontado los modelos DSGE con los datos, esto, por supuesto, ha significado un gran reto en la investigación macroeconómica. [Plosser \(1989\)](#) utiliza dos formas de confrontar un modelo RBC con los datos: momentos estadísticos y tasas de crecimiento centrados en cero.

En este documento intentamos contrastar un modelo DSGE con los datos. Suponemos que los *shocks* a la PTF son las únicas fuerzas que generan las fluctuaciones macroeconómicas. Para lograr este objetivo usamos datos de Estados Unidos desde el primer trimestre de 1950 hasta el tercer trimestre de 2019. Usamos tres técnicas. Primero, se resuelve el modelo, se simula, se aplica un filtro estadístico y se obtienen los momentos estadísticos del modelo, los cuales se comparan con los momentos obtenidos de los datos. Segundo, se usan los residuales del proceso que sigue la PTF para simular los ciclos del modelo, estos también se comparan con los ciclos obtenidos de los datos. Tercero, se comparan las impulso-respuesta del modelo DSGE con las de un modelo SVAR. Como un ejercicio adicional, se estima el modelo usando métodos bayesianos.

El resto del documento está organizado de la siguiente manera. En la sección 2 presentamos la estructura del modelo, en la sección 3 reportamos los resultados, en la sección 4 se estima el modelo y, finalmente, concluimos el trabajo en la sección 5.

2 El modelo

Consideramos un modelo estocástico de equilibrio general con progreso técnico aumentativo de trabajo, X_t , para una economía cerrada y sin gobierno. Siendo $X_t = \gamma_x X_{t-1}$, con $\gamma_x = 1 + g_t$ la tasa de crecimiento de la productividad laboral. En esta economía existe un planeador central que maximiza el valor esperado de la función de utilidad, $\mathbb{E}_t u(C_t, 1 - h_t)$, que depende del consumo de bienes, C_t , y el tiempo dedicado al ocio, $1 - h_t$, siendo h_t las horas de trabajo. El supuesto del agente representativo implica

la inexistencia de mercados, de manera que no existen precios ni de los bienes ni de los factores. Suponemos una función de utilidad Cobb-Douglas anidada en una CRRA. El problema para un horizonte infinito viene dado por:

$$\max_{\{C_t, h_t\}_{t=0}^{\infty}} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \frac{[C_t^\gamma (1 - h_t)^{1-\gamma}]^{1-\sigma} - 1}{1 - \sigma}$$

sujeto a las restricciones:

$$\begin{aligned} Y_t &= C_t + I_t \\ K_{t+1} &= I_t + (1 - \delta)K_t \\ Y_t &= A_t K_t^\alpha (X_t h_t)^{1-\alpha} \\ \ln A_{t+1} &= (1 - \rho) \ln A_{ss} + \rho \ln A_t + \varepsilon_{t+1} \end{aligned}$$

Siendo $\beta \in (0, 1)$ el factor de descuento, $\gamma \in (0, 1)$ la participación del consumo en la utilidad, $\sigma > 1$ la aversión al riesgo intertemporal, $\alpha \in (0, 1)$ la participación del factor capital en la producción, $\delta \in (0, 1)$ la tasa de depreciación del capital y $\rho \in (0, 1)$ la persistencia del *shock* tecnológico, $\varepsilon_t, \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Así, el conjunto de parámetros viene dada por; $\theta = \{\beta, \gamma, \sigma, \alpha, \delta, \gamma_x, \rho, \sigma_\varepsilon\}$.

Es claro que en estado estable las variables crecen al ritmo de X_t , excepto, probablemente las horas de trabajo. Esto sugiere normalizar las variables entre X_t , es decir $Y_t/X_t, C_t/X_t, I_t/X_t$ y K_t/X_t . Estas nuevas variables se escriben en letras minúsculas, con ellas se obtienen las condiciones de equilibrio y se escribe el sistema log-lineal con expectativas racionales, $\mathbb{E}_t f(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{z}_t) = 0$ y $g(\mathbf{z}_{t+1}, \mathbf{z}_t, \varepsilon_{t+1}) = 0$. Donde $\mathbf{x}_t = \{\hat{c}_t, \hat{h}_t, \hat{k}_{t+1}, \hat{i}_t, \hat{y}_t, \hat{A}_t\}$ es el conjunto de variables endógenas como desviaciones logarítmicas respecto al estado estable. Por ejemplo $\hat{c}_t = \log(c_t) - \log(c_{ss})$ es la desviación del consumo respecto al estado estacionario. Además, $\mathbf{z}_t = \{\varepsilon_t\}$ y ε_t es una innovación tecnológica. Para resolver este sistema existen muchos métodos, ver por ejemplo [Blanchard and Kahn \(1980\)](#), [King et al. \(1988a\)](#), [Uhlig \(1995\)](#), [Binder and Pesaran \(1997\)](#), [Klein \(1997\)](#), [Klein \(2000\)](#), [Sims \(2002\)](#) y [Christiano \(2002\)](#). Nosotros usamos el método [King et al. \(1988a\)](#) (Ver Anexo). El resultado que se obtiene son las funciones de política. En general, escribimos de la siguiente manera.

$$\mathbf{x}_t = \Gamma_x(\theta) \mathbf{x}_{t-1} + \Gamma_\varepsilon(\theta) \varepsilon_t \quad (1)$$

3 Resultados

Para simular numéricamente el modelo usamos los siguientes valores de los parámetros; $\alpha = 0,36$; $\delta = 0,025$; $\gamma = 0,8$; $\sigma = 2$ y $\gamma_x = 1,04$. En el caso del factor de descuento se obtiene a partir de $\tilde{\beta} = \beta \gamma_x^{\gamma(1-\sigma)}$, siendo $\tilde{\beta}$ el factor de descuento obtenido luego

de normalizar el modelo, por lo tanto, para un $\beta = 0,99$ se tiene que; $\tilde{\beta} = 0,96$. Usando el componente cíclico de la PTF de Estados Unidos se obtiene que $\rho = 0.76$ y $\sigma_\varepsilon = 0.006998$. Con este conjunto de valores podemos obtener el estado estacionario del modelo que vienen dados por; $c_{ss} = 1,1769$; $h_{ss} = 0,7653$; $k_{ss} = 4,9501$; $i_{ss} = 0,3218$ y $y_{ss} = 1,4986$. Por definición se tiene que $\varepsilon_{ss} = 0$ y por lo tanto $A_{ss} = 1$.

Organizando las variables de lo más exógeno o lo más endógeno, las matrices $\Gamma_x(\theta)$ y $\Gamma_\varepsilon(\theta)$ vienen dadas por:

$$\mathbf{x}_t = \begin{bmatrix} \widehat{A}_t \\ \widehat{k}_t \\ \widehat{h}_t \\ \widehat{y}_t \\ \widehat{i}_t \\ \widehat{c}_t \end{bmatrix} ; \quad \Gamma_x(\theta) = \begin{bmatrix} 0 & 0,7609 \\ 0,9316 & 0,1578 \\ -0,0227 & 0,0879 \\ 0,3455 & 0,8172 \\ -0,0941 & 2,5214 \\ 0,4385 & 0,4566 \end{bmatrix} ; \quad \Gamma_\varepsilon(\theta) = \begin{bmatrix} 1,0000 \\ 0,2070 \\ 0,1155 \\ 1,0739 \\ 3,3135 \\ 0,6000 \end{bmatrix}$$

En la [Figura 1](#) se observan las respuestas de las variables endógenas ante una innovación tecnológica que afecta positivamente a la productividad total de los factores, la cual, a su vez, genera efectos positivos sobre los factores productivos, capital y trabajo. La historia es que la producción también aumenta positivamente, hasta un 0,8% respecto a su estado estable. Para reestablecer el equilibrio, los factores de demanda, el consumo y la inversión, también se desvían positivamente hasta un 0,48% y 2,4% respectivamente. Los efectos de los choques de tecnología desaparecen en el tiempo y el resto de las variables, como consecuencia, convergen al equilibrio inicial.

Ahora bien, la cuestión que surge es ¿Qué tanto reproduce este modelo los hechos estilizados de los ciclos? ¿Qué medidas se pueden usar para contrastar el modelo con los datos?. Para responder a estas preguntas usamos tres métodos. Primero, comparamos los momentos estadísticos del modelos con los momentos obtenidos de los datos. Segundo, efectuamos una simulación de los ciclos económicos y, tercero, comparamos las funciones de impulso-respuesta del modelo con las de un modelo de Vectores Autorregresivos Estructural (SVAR).

En el primer caso, se efectúa una simulación estocástica, se filtra usando el filtro Hodrick-Prescott ($\lambda = 1600$), se obtienen los momentos estadísticos y se comparan con los datos. La [Tabla 1](#) muestran estos resultados para las variables claves del modelo. Se observa que la desviación estándar del producto estimado, del 1,14%, son muy próximos a los datos, del 1,52%. En el caso del resto de las variables el modelo subestima la desviación estándar. Por otro lado, los coeficientes de correlación del modelo sobreestiman en comparación con las correlaciones de los datos. Finalmente, los coeficientes de persistencia del modelo están muy próximos a los datos, sólo en el caso del producto y consumo. En el caso de la inversión y las horas de trabajo son subestimados.

A continuación, para simular los ciclos, construimos la PTF aproximando mediante

el residuo de Solow, para lo cual usamos la función de producción Cobb-Douglas. A este residuo lo extraemos el componente cíclico mediante HP, con el cual efectuamos una regresión mínimo cuadrática y obtenemos los residuos. Usamos estos residuos para efectuar la simulación de los ciclos de las variables claves. En la [Figura 2](#) se observa el comportamiento de los ciclos del producto obtenidos a partir del modelo en comparación con los datos. Se puede observar que el modelo es capaz de reproducir la trayectoria en cuanto a las fases de expansión y las fases de contracción, así como los máximos y mínimos del ciclo. En general, el modelo reproduce el comportamiento del ciclo del producto durante todo el periodo. La correlación entre ambas series (observado y estimado) es del 86%. En cuanto al consumo y la inversión se observa que el modelo también es capaz de replicar el comportamiento cíclico. Se obtienen unas correlaciones de 70% y 75%, respectivamente, que son bastante altas. Sin embargo, el modelo no reproduce muy bien las horas de trabajo, de hecho el modelo subestima el comportamiento cíclico. Bajo este contexto, la correlación que se encuentra es del 43%.

Finalmente, formulamos un modelo SVAR, de orden 1, con la finalidad de comprar las respuesta de las variables ante choques de tecnología con las del modelo DSGE. El modelo propuesta es el siguiente:

$$\mathbf{A}_0 \mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{B} \mathbf{u}_t \quad (2)$$

Donde el vector \mathbf{y}_t contiene las siguientes variables; la PTF ($\hat{\mathbf{A}}$), horas de trabajo ($\hat{\mathbf{h}}$), producto ($\hat{\mathbf{y}}_t$), inversión ($\hat{\mathbf{i}}_t$) y consumo ($\hat{\mathbf{c}}_t$), en ese orden. Todas las variables vienen expresados en términos de desviaciones logarítmicas de su tendencia, para lo cual, se aplica el filtro HP ($\lambda = 1600$). El vector \mathbf{y}_t es de orden 5×1 , la matriz \mathbf{A}_0 , que captura las relaciones contemporáneas de las variables, es de orden 5×5 , y la matriz \mathbf{A}_1 , relaciona las variables endógenas con los rezagos de ellas mismas, también es de orden 5×5 .

Para la identificación del modelo SVAR se supone que el sistema es recursivo, esto implica que \mathbf{A}_0 es típicamente una matriz triangular inferior con unos en la diagonal principal. Básicamente, es una historia sobre una determinada variable endógena explicada por aquellas que están “arriba” en el sistema. Por naturaleza, la fuente de perturbaciones proviene de los choques a la PTF, lo que implica que esta variable es un proceso AR(1), para lo cual hacemos, por lo tanto, imponemos las restricciones a la primera fila de la matriz \mathbf{A}_1 haciendo todos ceros excepto el primer elemento, que captura la persistencia (\mathbf{a}_{11}), de modo que:

$$\hat{\mathbf{A}}_t = \mathbf{a}_{11} \hat{\mathbf{A}}_{t-1} + \boldsymbol{\epsilon}_t$$

Con los supuestos de identificación del modelo empírico, los resultados de los modelos (1) y (2) son comparables¹. En la [Figura 4](#) se observas las respuestas de las variables

¹Para el estimar el modelo SVAR primero se estima un modelo VAR reducido, para ello pre-multiplicamos el modelo estructural por \mathbf{A}_0^{-1} , de esta forma tenemos: $\mathbf{x}_t = \mathbf{A} \mathbf{x}_{t-1} + \boldsymbol{\epsilon}_t$. Donde $\mathbf{A} = \mathbf{A}_0^{-1} \mathbf{A}_1$; $\boldsymbol{\epsilon}_t = \mathbf{S} \mathbf{u}_t$; $\mathbf{S} = \mathbf{A}_0^{-1} \mathbf{B}$

ante choques de PTF. En primer lugar, como se mencionó anteriormente, los coeficientes ρ y σ_ε son comparables en en ambos modelos, DSGE y SVAR, de modo que el proceso que sigue \widehat{A}_t es el mismo en ambos casos. Con respecto al resto de las variables, para un intervalo de confianza del 95%, el modelo es capaz de replicar la dinámica de transición hasta el periodo 20, dado que la trayectoria se encuentra en las bandas de confianza. En el caso de la inversión también se encuentra este comportamiento excepto entre los periodos 14-18, en la que el modelo parece sobre estimar la reacción. En el caso del consumo, aunque predice bien la dirección del movimiento, esta más no se ajusta muy bien. Finalmente, en el caso de las horas de trabajo no es posible observar un buen ajuste, de hecho subestima la reacción hasta el periodo 10.

4 Estimación bayesiana

La estimación bayesiana de los modelos DSGE se volvieron populares a inicios de los 2000. Los trabajo que mayor influencia han genera son; [Schorfheide \(2000\)](#), [Lubik and Schorfheide \(2005\)](#), [Rabanal and Rubio-Ramírez \(2005\)](#), [Herbst and Schorfheide \(2015\)](#), entre otros. En general, la estimación bayesiana es un puente entre la calibración y la máxima verosimilitud La calibración se hereda a través de la especificación de *priors* y el enfoque de máxima verosimilitud entra a través del proceso de estimación basado en confrontar el modelo con los datos. Juntos, los *priors* pueden verse como ponderaciones en la función de probabilidad para dar más importancia a ciertas áreas del subespacio de parámetros. Estos dos bloques de construcción, *priors* y funciones de verosimilitud, están relacionados por la regla de Bayes.

Antes de ello es necesario escribir el modelo en su forma estado-espacio:

$$\begin{aligned} \mathbf{s}_t &= \Gamma_s(\theta)\mathbf{s}_{t-1} + \Gamma_\varepsilon(\theta)\boldsymbol{\varepsilon}_t \\ \mathbf{y}_t &= \Phi(\theta)\mathbf{s}_t + \mathbf{u}_t \end{aligned}$$

Donde \mathbf{s}_t es un vector de variables de estado, \mathbf{y}_t un vector de observaciones y \mathbf{u}_t es un error de medición. La ecuación de observación relaciona la variable \mathbf{s}_t con el vector de observaciones \mathbf{y}_t . A continuación, sea $\mathbf{Y} = \{\mathbf{y}_t\}_1^T$ el conjunto de observaciones. Los parámetros estructurales se resumen en θ . La función de verosimilitud $\mathcal{L}(\theta|\mathbf{Y})$ se combina con una densidad previa, “*prior*”, $p(\theta)$ para formar una densidad posterior $p(\theta|\mathbf{Y})$, usando el teorema de Bayes.

$$p(\theta|\mathbf{Y}) \propto \mathcal{L}(\theta|\mathbf{Y})p(\theta)$$

Donde \propto representa una proporcionalidad. Este teorema permite la actualización continua de la información sobre los parámetros de la distribución conforme se generan más observaciones. Luego se utiliza el filtro de Kalman, un método de predicción y corrección, para evaluar la función de verosimilitud. Finalmente se utiliza el algoritmo

Metrópolis-Hastings para estimar los parámetros a partir del muestreo de la distribución a posteriori.

Los priors de los parámetros se observan en la [Tabla 2](#). En la mayoría de los casos suponemos una distribución beta, excepto para la desviación estándar del error de la PTF, que suponemos sigue una distribución gamma inversa, dado que se trata de una desviación estándar. Usando el componente cíclico del producto como la variable observable, con 274 observaciones, podemos estimar los parámetros del modelo. En la [Tabla 2](#) se observan los resultados de las estimaciones y sus intervalos de confianza al 95%. Se efectúa 5000 réplicas de simulación para el algoritmo Metropolis-Hasting. Las distribuciones de las priors en comparación con las posteriors se observan en la [Figura 7](#), así mismo las funciones de impulso respuesta bayesiano se reportan en la [Figura 8](#) y en la [Figura 10](#) se reportan las optimizaciones de los estimadores.

5 Conclusiones

En este documento hemos confrontado un modelo DSGE simple con los datos. En general, el modelo reproduce la dirección del movimiento de las variables claves; producto, consumo e inversión, pero en cuanto a las magnitudes se observan algunas diferencias, aunque estas no parecen ser significativas. Sin embargo, el modelo no es capaz de reproducir el comportamiento de las horas de trabajo. Este punto ha generado una gran cantidad de literatura y ha sido motivo para desarrollar otros modelos que incorporen imperfecciones en el mercado laboral. Una recomendación es que sería interesante explorar los efectos del mercado laboral formal e informal. Esto es particularmente importante para las economías emergentes, dada la estructura del mercado laboral.

References

- Binder, M. and Pesaran, M. H. (1997). Multivariate linear rational expectations models: characterization of the nature of the solutions and their fully recursive computation. *Econometric Theory*, 13(6):877–888.
- Blanchard, O. J. and Kahn, C. M. (1980). The solution of linear difference models under rational expectations. *Econometrica: Journal of the Econometric Society*, 48(5):1305–1311.
- Christiano, L. J. (2002). Solving dynamic equilibrium models by a method of undetermined coefficients. *Computational Economics*, 20(1-2):21–55.
- Herbst, E. P. and Schorfheide, F. (2015). *Bayesian estimation of DSGE models*. Princeton University Press.
- King, R. G., Plosser, C. I., and Rebelo, S. T. (1988a). Production, growth and business cycles: I. the basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195–232.
- King, R. G., Plosser, C. I., and Rebelo, S. T. (1988b). Production, growth and business cycles: II. new directions. *Journal of Monetary Economics*, 21(2-3):309–341.
- Klein, P. (1997). Using the generalized schur form to solve a system of linear expectational difference equations. *manuscript, IIES, Stockholm University*.
- Klein, P. (2000). Using the generalized schur form to solve a multivariate linear rational expectations model. *Journal of economic dynamics and control*, 24(10):1405–1423.
- Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.
- Long, J. B. and Plosser, C. I. (1983). Real business cycles. *Journal of Political Economy*, 91(1):39–69.
- Lubik, T. and Schorfheide, F. (2005). A bayesian look at new open economy macroeconomics. *NBER macroeconomics annual*, 20:313–366.
- Plosser, C. I. (1989). Understanding real business cycles. *The Journal of Economic Perspectives*, 3(3):51–77.
- Rabanal, P. and Rubio-Ramírez, J. F. (2005). Comparing new keynesian models of the business cycle: A bayesian approach. *Journal of Monetary Economics*, 52(6):1151–1166.
- Schorfheide, F. (2000). Loss function-based evaluation of dsge models. *Journal of Applied Econometrics*, 15(6):645–670.

- Sims, C. A. (2002). Solving linear rational expectations models. *Computational economics*, 20(1):1–20.
- Uhlig, H. F. (1995). A toolkit for analyzing nonlinear dynamic stochastic models easily. *Discussion Paper 01*.
- Wickens, M. (2012). *Macroeconomic Theory: A Dynamic General Equilibrium Approach*. Princeton University Press. 2da Edition.

A Appendix

Figura 1: Respuesta ante un choque de productividad total de factores

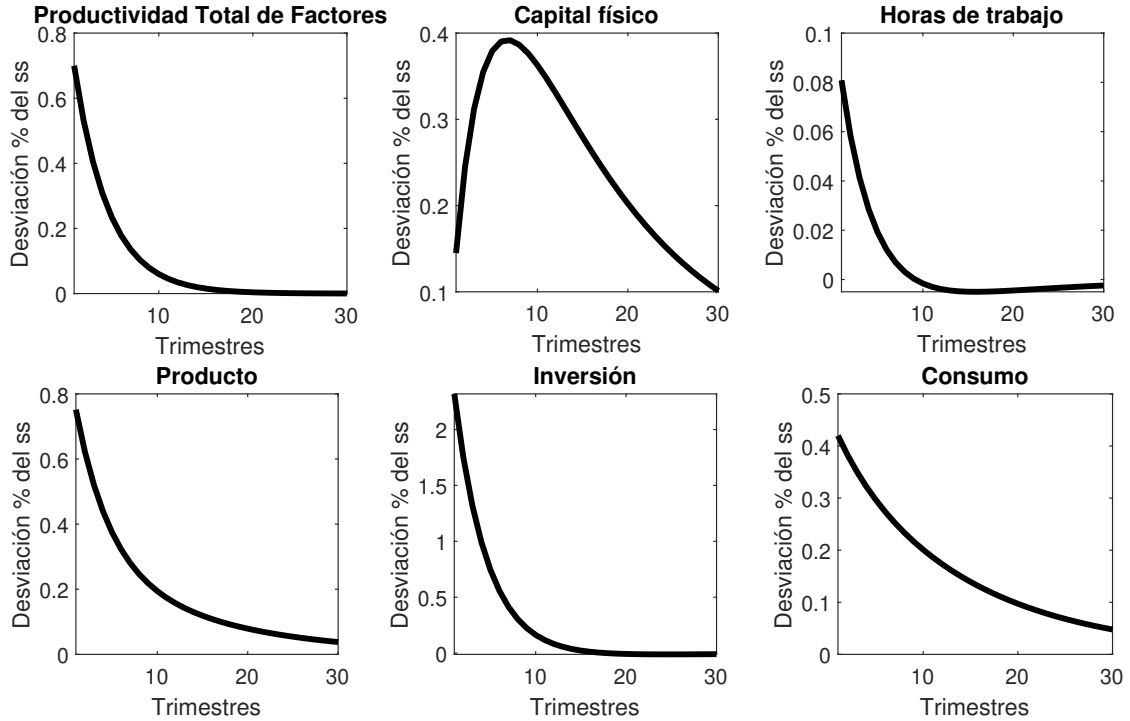


Table 1: Momentos estadísticos, modelo versus datos

Variables	$\sigma(\mathbf{x})$	$\frac{\sigma(\mathbf{x})}{\sigma(\hat{\mathbf{y}})}$	$\rho(\mathbf{x}_t, \hat{\mathbf{y}}_t)$	$\rho(\mathbf{x}_t, \mathbf{x}_{t-1})$
Datos				
Producto	1,52	1,00	1,00	0,83
Consumo	1,23	0,81	0,79	0,81
Inversión	6,73	4,44	0,86	0,81
Horas de trabajo	1,03	0,68	0,79	0,89
Modelo				
Producto	1,14	1,00	1,00	0,78
Consumo	0,78	0,68	0,97	0,86
Inversión	3,14	2,74	0,95	0,70
Horas de trabajo	0,11	0,09	0,89	0,68

Fuente: Elaboración propia. Nota, $\hat{\mathbf{y}}_t$ es la desviación logarítmica del producto y $\hat{\mathbf{x}}_t$ el resto de las variables (consumo, inversión y horas de trabajo).

Figura 2: Producto: modelo versus datos

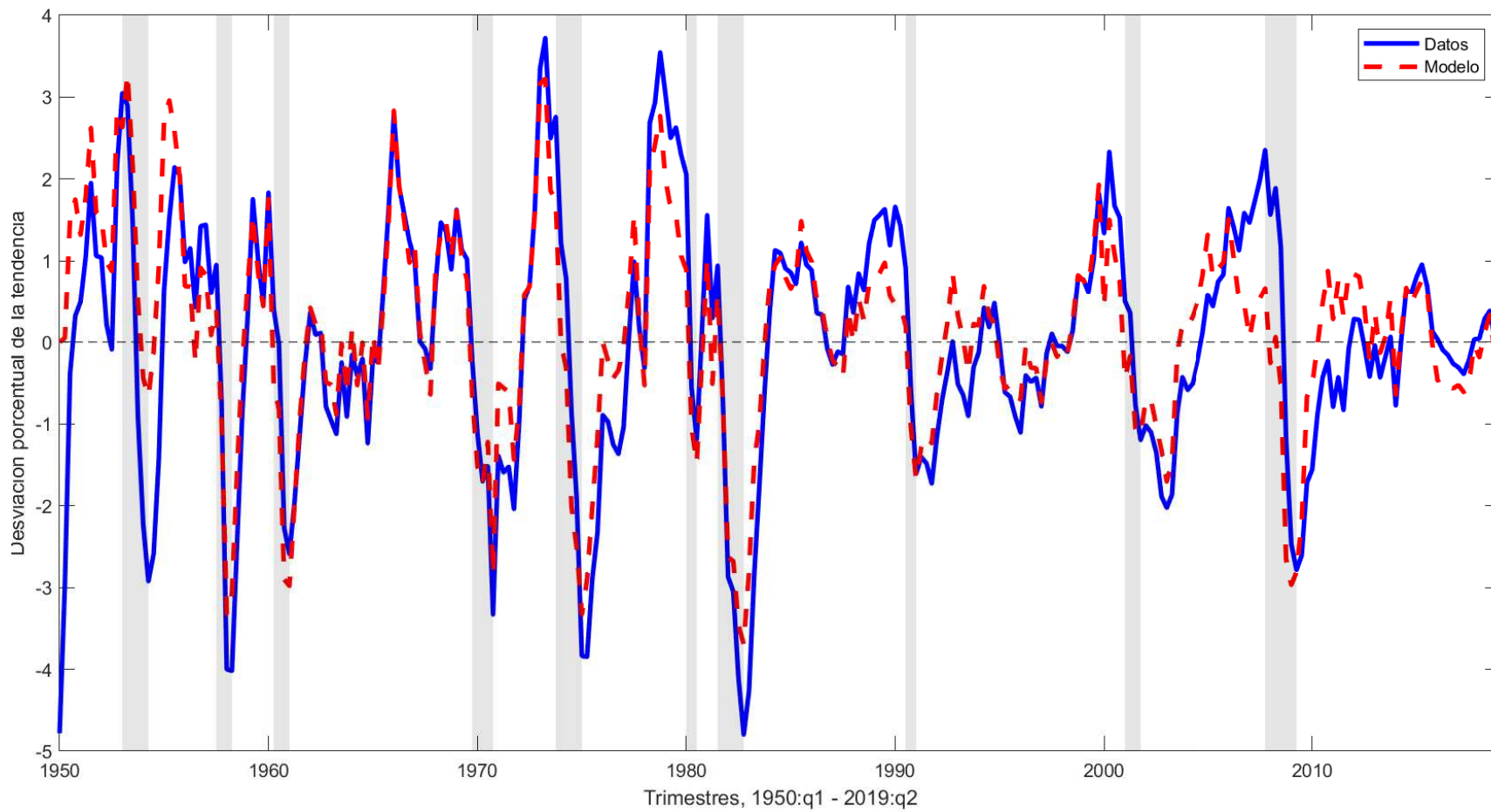


Figura 3: Consumo: modelo versus datos

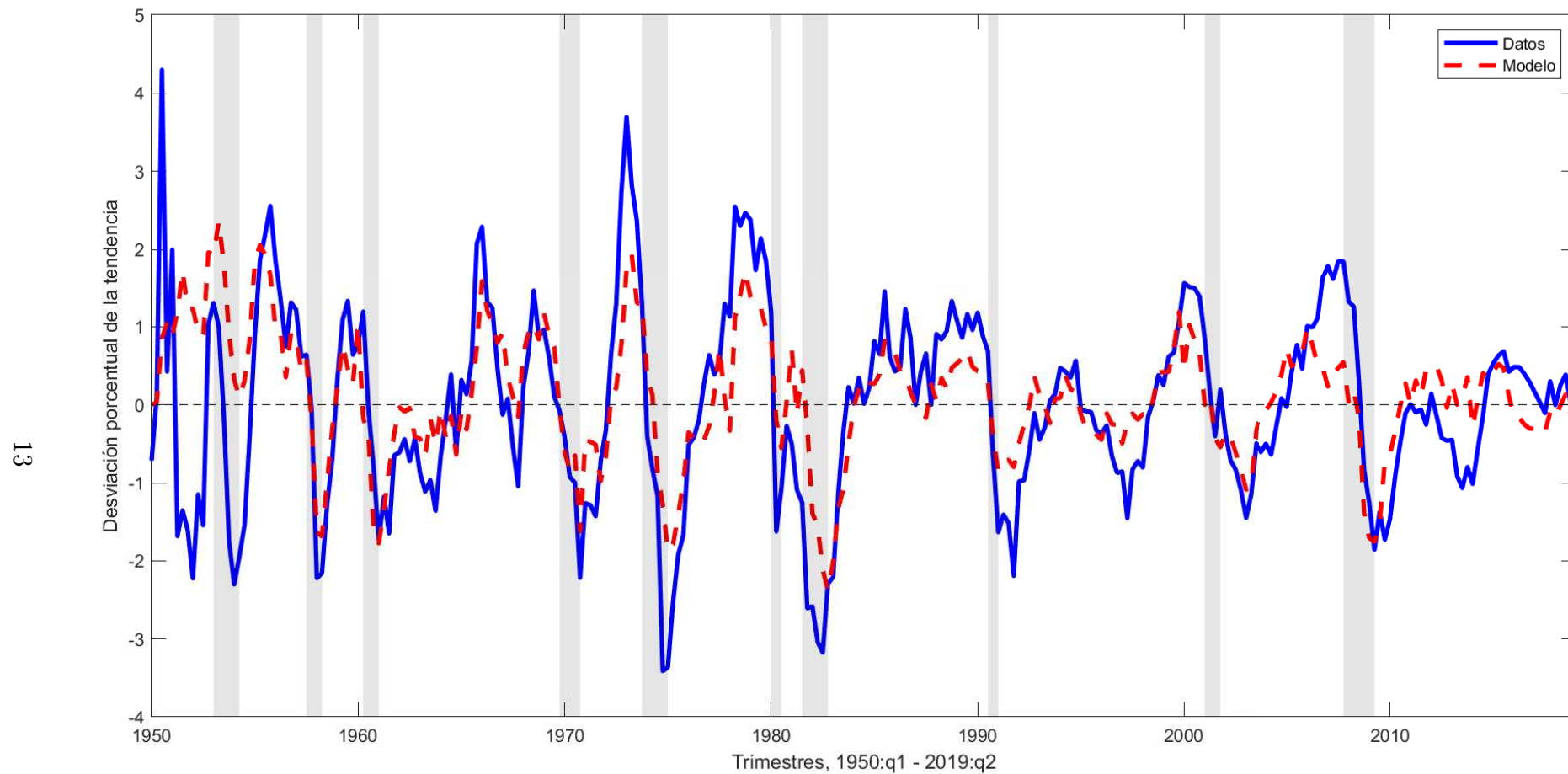


Figura 4: Inversión: modelo versus datos

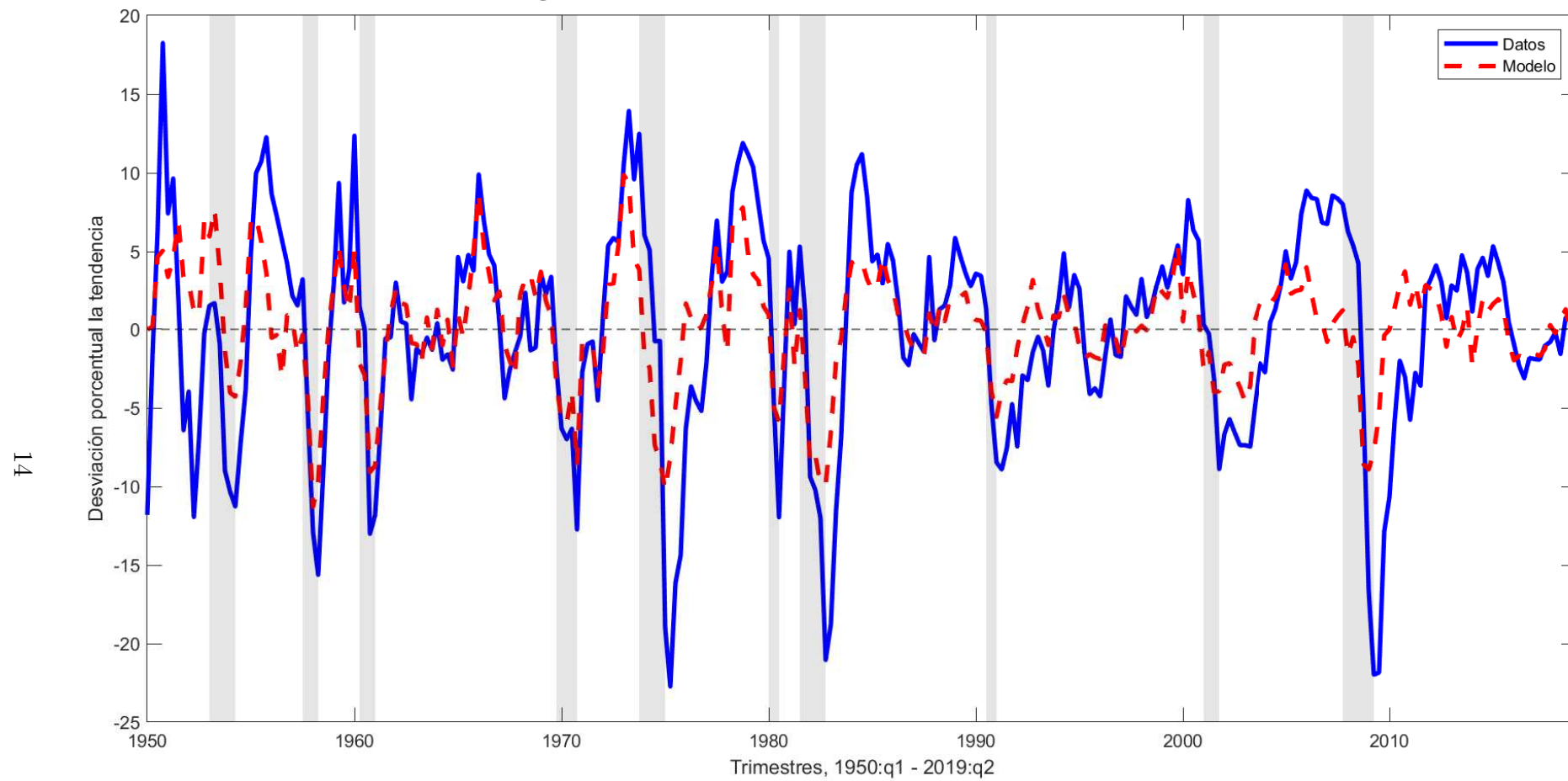


Figura 5: Horas de trabajo: modelo versus datos

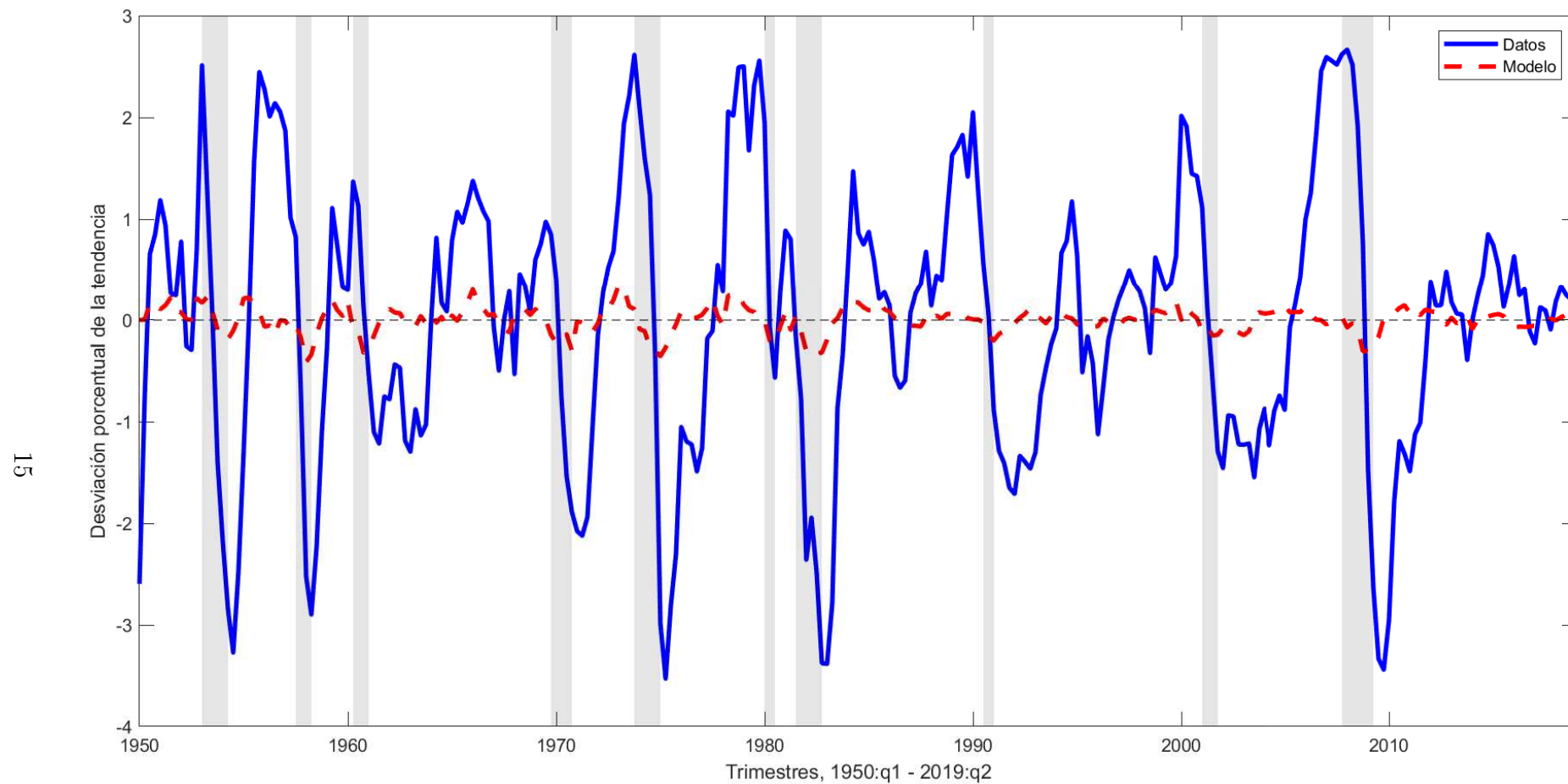


Figura 6: Respuesta ante un choque de productividad total de factores

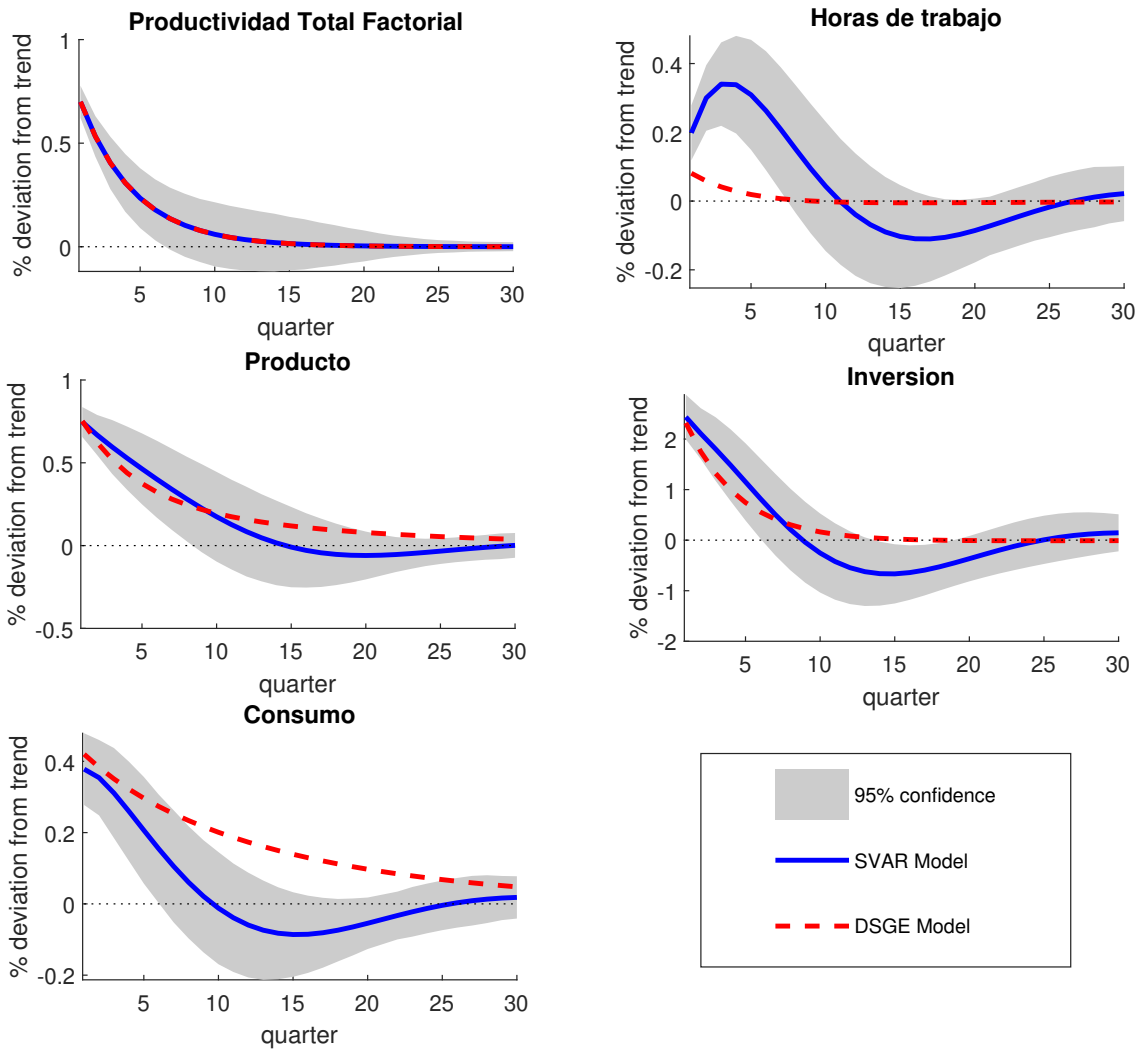


Table 2: Estimación bayesiana

Parámetros	Tipo de distribución	Prior	Posterior (Media)	Intervalo HDP 90%	Metropolis Hasting	Posterior Maximización
α	Beta	0,3560 (0,0200)	0,3542 (0,0200)	[0,3222 ; 0,3866]	0,0126	0,0126
β	Beta	0,9930 (0,0020)	0,9928 (0,0020)	[0,9897 ; 0,9963]	0,3542	0,3565
δ	Beta	0,0100 (0,0050)	0,0112 (0,0050)	[0,0034 ; 0,0215]	0,9928	0,9936
σ	Beta	0,6500 (0,0500)	0,6533 (0,0500)	[0,5709 ; 0,7325]	0,0112	0,0074
θ	Beta	0,3500 (0,0200)	0,3485 (0,0200)	[0,3172 ; 0,3801]	0,6533	0,6532
ρ	Beta	0,1290 (0,2230)	0,9684 (0,2230)	[0,9437 ; 0,9946]	0,3485	0,3500
σ_ε	Gamma inversa	0,0350 inf	0,0126 inf	[0,0111 ; 0,0138]	0,9684	0,9713
Densidad marginal de la Verosimilitud (Media armónica modificada)					550,0976	
Densidad marginal de la Verosimilitud (Aproximación de Laplace)					550,1793	

Entre paréntesis desviación estándar.

Figura 7: Priors y Posterior

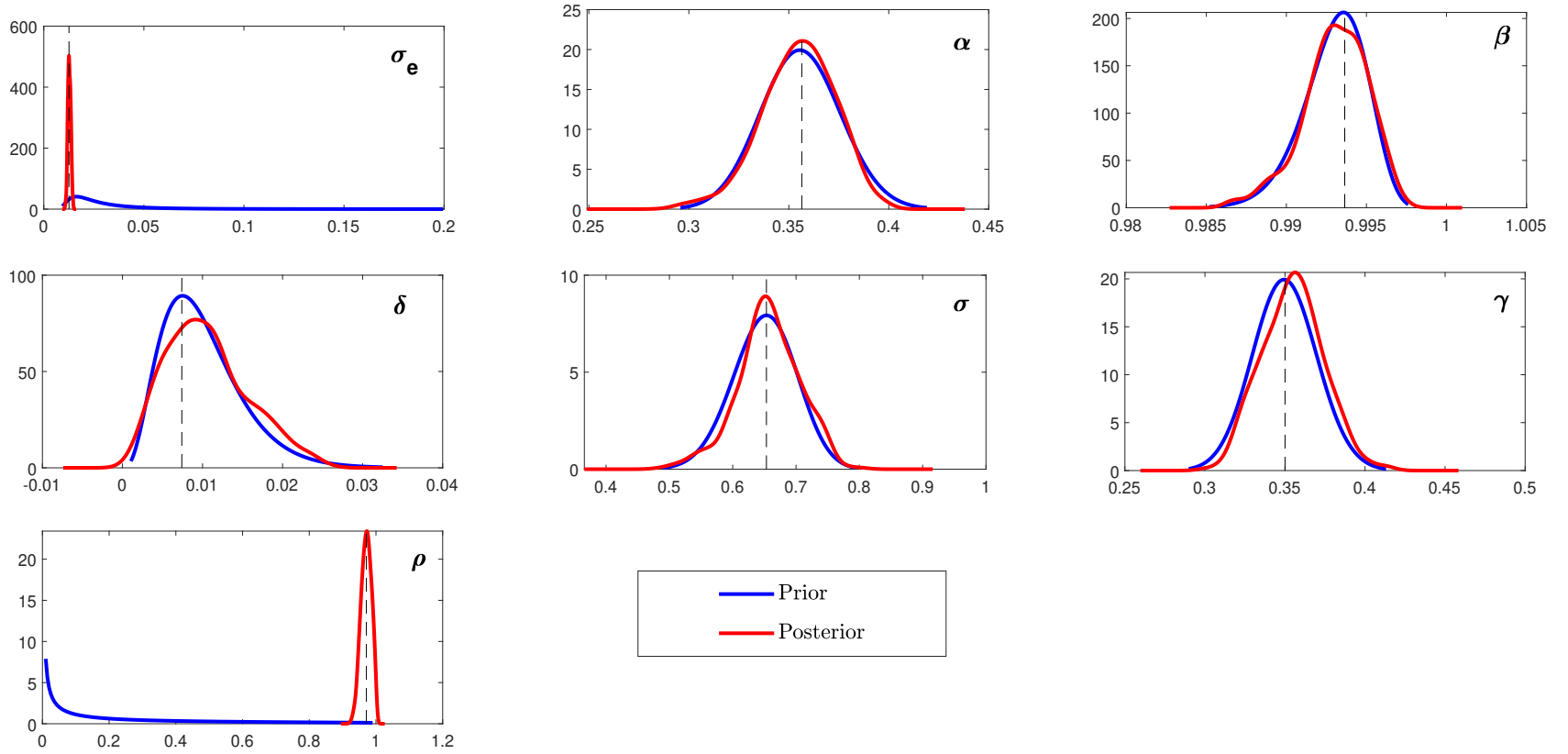


Figura 8: Funciones de impulso respuesta bayesiano

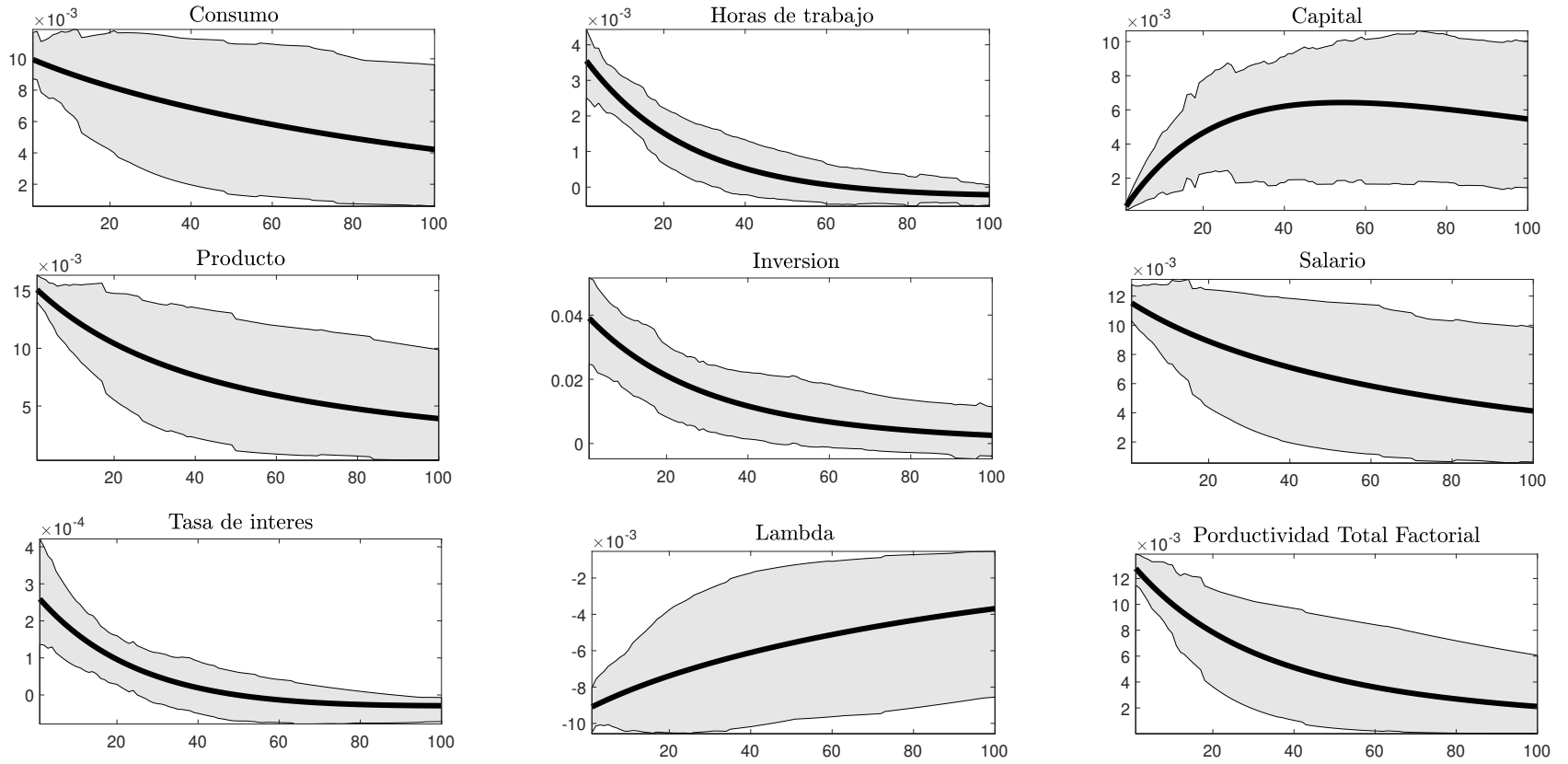


Figura 9: Descomposición histórica de la varianza

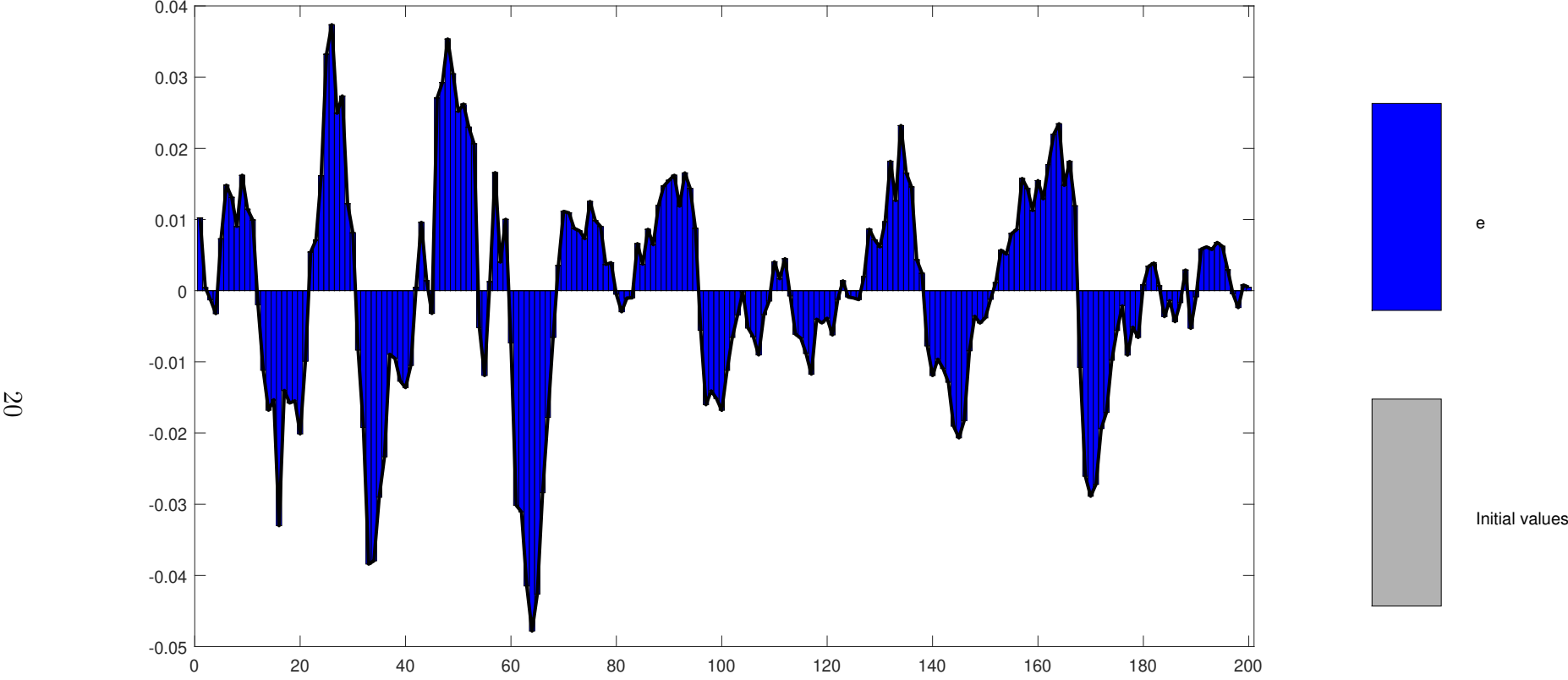


Figura 10: Optimización de los estimadores bayesianos

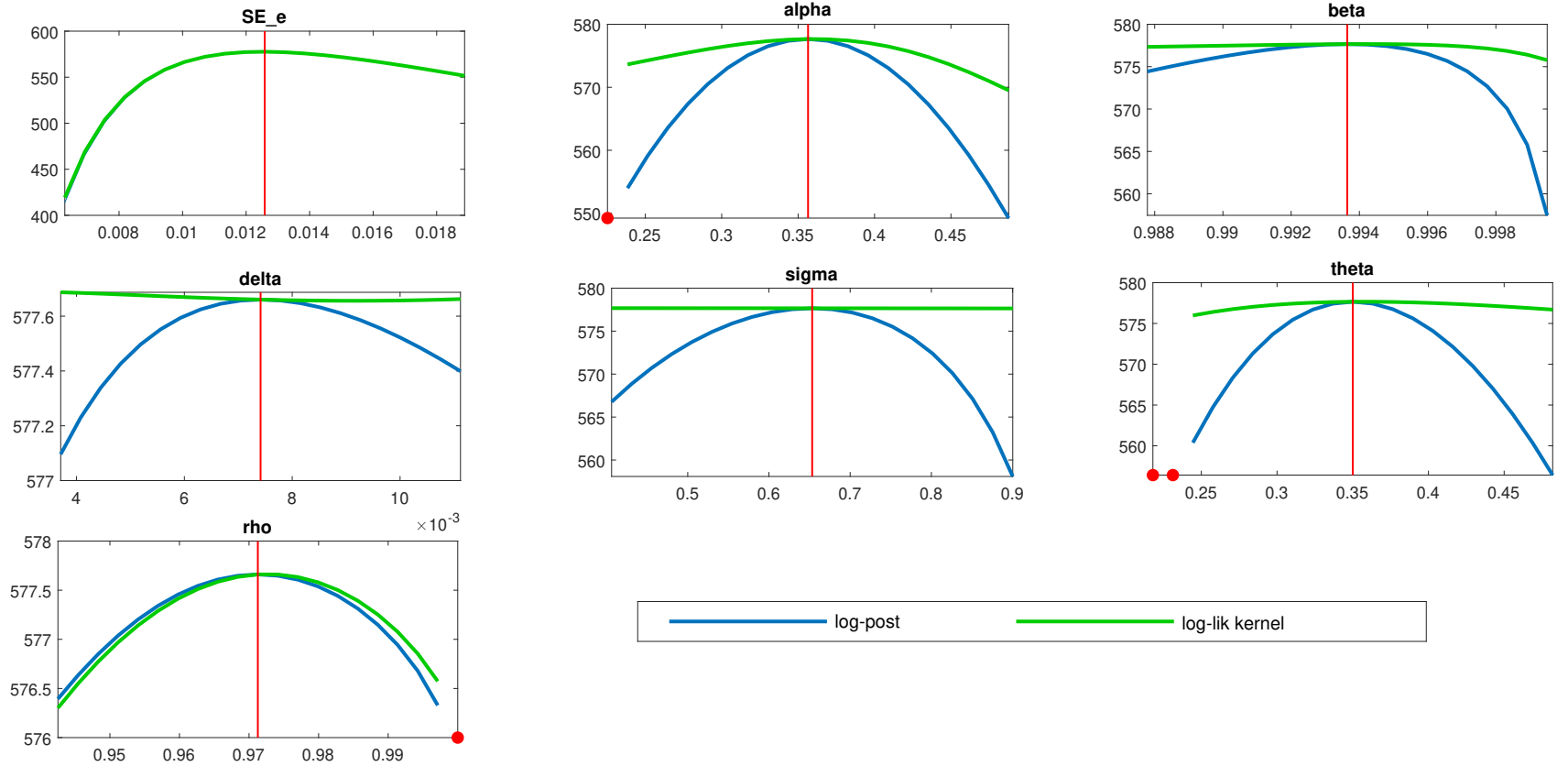


Figura 11: Diagnóstico de convergencia univariable MH

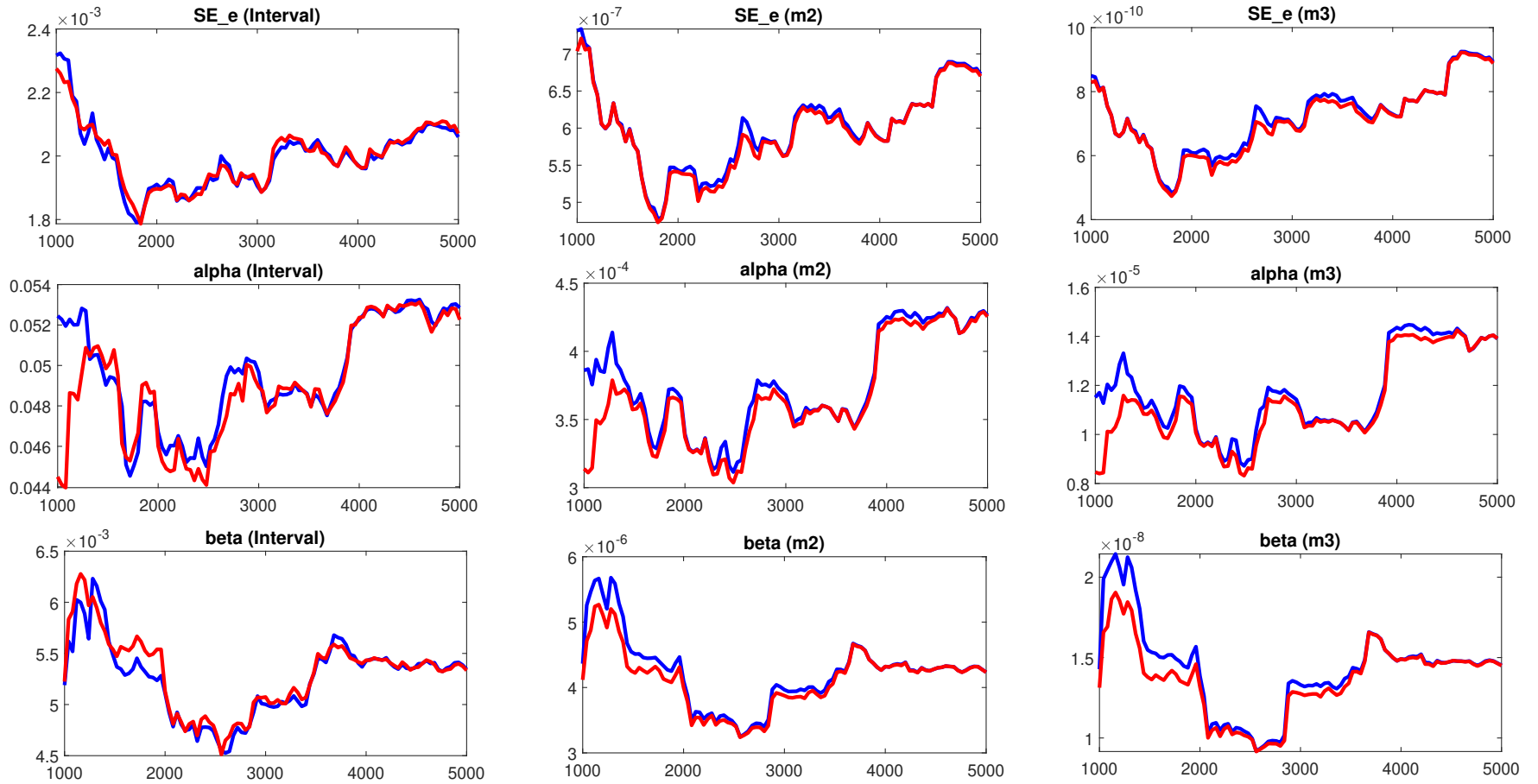
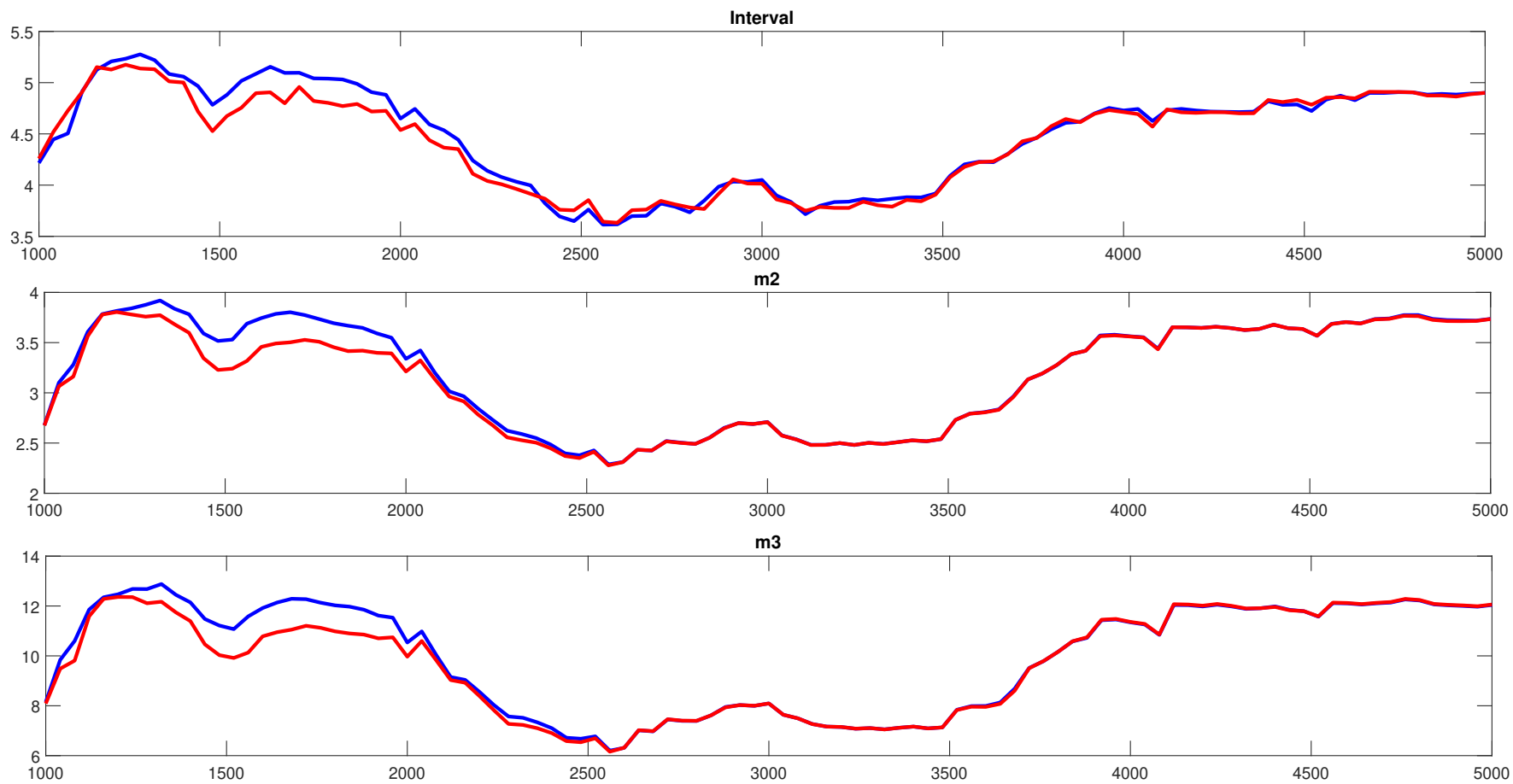


Figura 12: Diagnóstico de convergencia multivariable MH



B Appendix

Para resolver el modelo inicialmente se normaliza las variables. Seguimos el procedimiento estándar, ver por ejemplo [King et al. \(1988a\)](#) y [Wickens \(2012\)](#), crecimiento óptimo del capítulo 3. Definimos las siguientes variables como unidades de eficiencia.

$$y_t = \frac{Y_t}{X_t} \quad ; \quad c_t = \frac{C_t}{X_t} \quad ; \quad i_t = \frac{I_t}{X_t} \quad ; \quad k_t = \frac{K_t}{X_t}$$

Usando estas definiciones, además de la solución para X_t , $X_t = X_0 \gamma_x^t$ con $X_0 = 1$, el modelo normalizado queda de la siguiente manera:

$$\max_{\{c_t, h_t\}_{t=0}^{\infty}} \mathbb{E}_t \sum_{t=0}^{\infty} \tilde{\beta}^t \frac{[c_t^\gamma (1 - h_t)^{1-\gamma}]^{1-\sigma} - \gamma_x^{-\gamma(1-\sigma)t}}{1 - \sigma}$$

Con las siguientes restricciones

$$\begin{aligned} \mathcal{A}_t k_t^\alpha h_t^{1-\alpha} &= c_t + \gamma_x k_{t+1} - (1 - \delta)k_t \\ \ln \mathcal{A}_{t+1} &= (1 - \rho) \ln \mathcal{A}_{ss} + \rho \ln \mathcal{A}_t + \varepsilon_{t+1} \end{aligned}$$

Donde $\tilde{\beta} = \beta \gamma_x^{\gamma(1-\sigma)}$

Usando la función de Lagrange, las condiciones de equilibrio del modelo y el proceso exógeno que sigue \mathcal{A}_t forman un sistema dinámico no lineales con expectativas racionales:

$$\gamma c_t^{(1-\sigma)\gamma-1} (1 - h_t)^{(1-\gamma)(1-\sigma)} - \lambda_t = 0 \quad (3)$$

$$-(1 - \gamma) c_t^{(1-\sigma)\gamma} (1 - h_t)^{-\sigma-(1-\sigma)\gamma} + (1 - \alpha) \lambda_t \mathcal{A}_t k_t^\alpha h_t^{-\alpha} = 0 \quad (4)$$

$$\tilde{\beta} \mathbb{E}_t [\lambda_{t+1} (\alpha \mathcal{A}_{t+1} k_{t+1}^{\alpha-1} h_{t+1}^{1-\alpha} + 1 - \delta)] - \gamma_x \lambda_t = 0 \quad (5)$$

$$c_t - \mathcal{A}_t k_t^\alpha h_t^{1-\alpha} + \gamma_x \mathbb{E}_t k_{t+1} - (1 - \delta)k_t = 0 \quad (6)$$

$$\ln \mathcal{A}_{t+1} - (1 - \rho) \ln \mathcal{A}_{ss} - \rho \ln \mathcal{A}_t - \varepsilon_{t+1} = 0 \quad (7)$$

Donde λ_t el multiplicador de Lagrange. El estado estacionario de este modelo viene dado por:

$$\begin{aligned}
h_{ss} &= \frac{1}{\Omega + 1} \\
k_{ss} &= \frac{\Theta^{\frac{1}{\alpha-1}}}{\Omega + 1} \\
c_{ss} &= \frac{(\Theta - \gamma_x + 1 - \delta)\Theta^{\frac{1}{\alpha-1}}}{\Omega + 1} \\
\lambda_{ss} &= \gamma \left[\frac{(\Theta - \gamma_x + 1 - \delta)\Theta^{\frac{1}{\alpha-1}}}{\Omega + 1} \right]^{(1-\sigma)\gamma-1} \left[\frac{\Omega}{\Omega + 1} \right]^{(1-\gamma)(1-\sigma)} \\
\Omega &= \frac{(1-\gamma)(\Theta - \gamma_x + 1 - \delta)}{\gamma(1-\alpha)\Theta} \\
\Theta &= \frac{\gamma_x - (1-\delta)\beta}{\alpha\beta}
\end{aligned}$$

A continuación, log-linealizamos las condiciones de equilibrio (3)-(7). Para lograr esto escribimos de la siguiente manera;

$$\mathbb{E}_t F(e^{\ln \mathbf{x}_t}, e^{\ln \mathbf{x}_{t+1}}) = 0$$

En donde $\mathbf{x}_t = \{c_t, h_t, \lambda_t, A_t, k_t\}$ y $\mathbf{x}_{t+1} = \{h_{t+1}, \lambda_{t+1}, A_{t+1}, k_{t+1}\}$. Aplicando la expansión de Taylor de primer orden:

$$\mathbb{E}_t F(e^{\ln \mathbf{x}_t}, e^{\ln \mathbf{x}_{t+1}}) = F(\cdot)|_{ss} + \frac{\partial F(\cdot)}{\partial \ln \mathbf{x}_t} \Big|_{ss} \widehat{\mathbf{x}}_t + \mathbb{E}_t \left[\frac{\partial F(\cdot)}{\partial \ln \mathbf{x}_{t+1}} \Big|_{ss} \widehat{\mathbf{x}}_{t+1} \right]$$

siendo $\widehat{\mathbf{x}}_t = \ln \mathbf{x}_t - \ln \mathbf{x}_{ss}$, obtenemos el siguiente sistema log-lineal:

$$\beta_0 \widehat{c}_t + \beta_1 \widehat{h}_t + \beta_2 \widehat{\lambda}_t = 0 \quad (8)$$

$$\beta_3 \widehat{c}_t + \beta_9 \widehat{h}_t + \beta_5 \widehat{\lambda}_t + \beta_6 \widehat{A}_t + \beta_7 \widehat{k}_t = 0 \quad (9)$$

$$\mathbb{E}_t \left(\beta_{10} \widehat{\lambda}_{t+1} + \beta_{11} \widehat{A}_{t+1} + \beta_{12} \widehat{k}_{t+1} + \beta_{13} \widehat{h}_{t+1} \right) + \beta_{14} \widehat{\lambda}_t = 0 \quad (10)$$

$$\beta_{15} \widehat{c}_t + \beta_{16} \widehat{A}_t + \beta_{17} \widehat{k}_t + \beta_{18} \widehat{h}_t + \beta_{19} \mathbb{E}_t \widehat{k}_{t+1} = 0 \quad (11)$$

$$\widehat{A}_{t+1} - \rho \widehat{A}_t - \varepsilon_{t+1} = 0 \quad (12)$$

Donde β_j , $j = 1, \dots, 9$, son funciones de los parámetros estructurales y del estado estable del modelo. En concreto se tienen:

$$\beta_0 = \gamma(\gamma - \sigma\gamma - 1)c_{ss}^{(1-\sigma)\gamma-1}(1 - h_{ss})^{(1-\gamma)(1-\sigma)}$$

$$\beta_1 = -\gamma(1-\gamma)(1-\sigma)c_{ss}^{(1-\sigma)\gamma-1}(1 - h_{ss})^{(1-\gamma)(1-\sigma)-1}h_{ss}$$

$$\beta_2 = -\gamma c_{ss}^{(1-\sigma)\gamma-1}(1 - h_{ss})^{(1-\gamma)(1-\sigma)}$$

$$\begin{aligned}
\beta_3 &= -(1-\gamma)(1-\sigma)\gamma c_{ss}^{(1-\sigma)\gamma} (1-h_{ss})^{-\sigma-(1-\sigma)\gamma} \\
\beta_4 &= (1-\gamma)(-\sigma-\gamma+\sigma\gamma) c_{ss}^{(1-\sigma)\gamma} (1-h_{ss})^{-\sigma-(1-\sigma)\gamma-1} h_{ss} \\
\beta_5 &= (1-\alpha) A_{ss} k_{ss}^\alpha h_{ss}^{-\alpha} \lambda_{ss} \\
\beta_6 &= (1-\alpha) \lambda_{ss} k_{ss}^\alpha h_{ss}^{-\alpha} A_{ss} \\
\beta_7 &= \alpha(1-\alpha) \lambda_{ss} A_{ss} k_{ss}^\alpha h_{ss}^{-\alpha} \\
\beta_8 &= -\alpha(1-\alpha) \lambda_{ss} A_{ss} k_{ss}^\alpha h_{ss}^{-\alpha} \\
\beta_9 &= (1-\gamma)(-\sigma-\gamma+\sigma\gamma) c_{ss}^{(1-\sigma)\gamma} (1-h_{ss})^{-\sigma-(1-\sigma)\gamma-1} h_{ss} - \alpha(1-\alpha) \lambda_{ss} A_{ss} k_{ss}^\alpha h_{ss}^{-\alpha} \\
\beta_{10} &= \alpha\beta A_{ss} k_{ss}^{\alpha-1} h_{ss}^{1-\alpha} \lambda_{ss} + \beta(1-\delta) \lambda_{ss} \\
\beta_{11} &= \alpha\beta \lambda_{ss} k_{ss}^{\alpha-1} h_{ss}^{1-\alpha} A_{ss} \\
\beta_{12} &= \alpha\beta(\alpha-1) \lambda_{ss} A_{ss} k_{ss}^{\alpha-1} h_{ss}^{1-\alpha} \\
\beta_{13} &= \alpha\beta(1-\alpha) \lambda_{ss} A_{ss} k_{ss}^{\alpha-1} h_{ss}^{1-\alpha} \\
\beta_{14} &= -\gamma_x \lambda_{ss} \\
\beta_{15} &= c_{ss} \\
\beta_{16} &= -k_{ss}^\alpha h_{ss}^{1-\alpha} A_{ss} \\
\beta_{17} &= -\alpha A_{ss} k_{ss}^\alpha h_{ss}^{1-\alpha} - (1-\delta) k_{ss} \\
\beta_{18} &= -(1-\alpha) A_{ss} k_{ss}^\alpha h_{ss}^{1-\alpha} \\
\beta_{19} &= \gamma_x k_{ss}
\end{aligned}$$

Ahora escribimos en términos matriciales. De (8) y (9) despejamos el vector de variables de control.

$$\mathbf{M}_{cc} \begin{bmatrix} \widehat{c}_t \\ \widehat{h}_t \end{bmatrix} = \mathbf{M}_{cs} \begin{bmatrix} \widehat{k}_t \\ \widehat{\lambda}_t \end{bmatrix} + \mathbf{M}_{ce} \widehat{A}_t \quad \Rightarrow \quad \begin{bmatrix} \widehat{c}_t \\ \widehat{h}_t \end{bmatrix} = \mathbf{M}_{cc}^{-1} \mathbf{M}_{cs} \begin{bmatrix} \widehat{k}_t \\ \widehat{\lambda}_t \end{bmatrix} + \mathbf{M}_{cc}^{-1} \mathbf{M}_{ce} \widehat{A}_t \quad (13)$$

Donde \mathbf{M}_{cc} es una matriz control-control, \mathbf{M}_{cs} matriz control-estado y \mathbf{M}_{ce} control-exógeno. Por otro lado, de las ecuaciones (10) y (11)

$$\mathbf{M}_{ss}^0 \mathbb{E}_t \begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{\lambda}_{t+1} \end{bmatrix} + \mathbf{M}_{ss}^1 \begin{bmatrix} \widehat{k}_t \\ \widehat{\lambda}_t \end{bmatrix} = \mathbf{M}_{sc}^0 \mathbb{E}_t \begin{bmatrix} \widehat{c}_{t+1} \\ \widehat{h}_{t+1} \end{bmatrix} + \mathbf{M}_{sc}^1 \begin{bmatrix} \widehat{c}_t \\ \widehat{h}_t \end{bmatrix} + \mathbf{M}_{se}^0 \mathbb{E}_t \widehat{A}_{t+1} + \mathbf{M}_{se}^1 \widehat{A}_t \quad (14)$$

Sustituyendo (13) en (14) obtenemos la ecuación fundamental del modelo.

$$\mathbb{E}_t \begin{bmatrix} \widehat{k}_{t+1} \\ \widehat{\lambda}_{t+1} \end{bmatrix} = \mathbf{W} \begin{bmatrix} \widehat{k}_t \\ \widehat{\lambda}_t \end{bmatrix} + \mathbf{R}\mathbb{E}_t \widehat{A}_{t+1} + \mathbf{Q}\widehat{A}_t \quad (15)$$

Donde \mathbf{W} , \mathbf{R} y \mathbf{Q} son las matrices fundamentales del modelo. La matriz \mathbf{W} determina la estabilidad dinámica del modelo y, junto al resto de las matrices, determinan los coeficientes de política. Estas matrices vienen definidas por:

$$\begin{aligned} \mathbf{W} &= [\mathbf{M}_{ss}^0 - \mathbf{M}_{sc}^0 \mathbf{M}_{cc}^{-1} \mathbf{M}_{cs}]^{-1} [\mathbf{M}_{ss}^1 - \mathbf{M}_{sc}^1 \mathbf{M}_{cc}^{-1} \mathbf{M}_{cs}] \\ \mathbf{R} &= [\mathbf{M}_{ss}^0 - \mathbf{M}_{sc}^0 \mathbf{M}_{cc}^{-1} \mathbf{M}_{cs}]^{-1} [\mathbf{M}_{se}^0 + \mathbf{M}_{sc}^0 \mathbf{M}_{cc}^{-1} \mathbf{M}_{ce}] \\ \mathbf{Q} &= [\mathbf{M}_{ss}^0 - \mathbf{M}_{sc}^0 \mathbf{M}_{cc}^{-1} \mathbf{M}_{cs}]^{-1} [\mathbf{M}_{se}^1 + \mathbf{M}_{sc}^1 \mathbf{M}_{cc}^{-1} \mathbf{M}_{ce}] \end{aligned}$$

Realizadas las operaciones, estas matrices vienen dados por:

$$\mathbf{W} = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \quad ; \quad \mathbf{R} = \begin{bmatrix} 0 \\ R_{21} \end{bmatrix} \quad ; \quad \mathbf{Q} = \begin{bmatrix} Q_{11} \\ Q_{21} \end{bmatrix}$$

Donde los elementos de estas matrices están dados de la siguiente manera (Esto es fácilmente demostrable)

$$\begin{aligned} W_{11} &= \frac{-(\beta_{10} + \alpha_4 \beta_{13})(\beta_{17} + \alpha_1 \beta_{15} + \alpha_3 \beta_{18})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \quad ; \quad W_{12} = \frac{-(\beta_{10} + \alpha_4 \beta_{13})(\alpha_2 \beta_{15} + \alpha_4 \beta_{18})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \\ W_{21} &= \frac{(\beta_{12} + \alpha_3 \beta_{13})(\beta_{17} + \alpha_1 \beta_{15} + \alpha_3 \beta_{18})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \quad ; \quad W_{22} = \frac{-\beta_{14} \beta_{19} + (\beta_{12} + \alpha_3 \beta_{13})(\alpha_2 \beta_{15} + \alpha_4 \beta_{18})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \\ R_{21} &= \frac{-\beta_{19}(\beta_{11} + \alpha_4 \beta_{13})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \\ Q_{11} &= \frac{-(\beta_{10} + \alpha_4 \beta_{13})(\beta_{16} + \alpha_5 \beta_{15} + \alpha_6 \beta_{18})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \\ Q_{21} &= \frac{(\beta_{12} + \alpha_3 \beta_{13})(\beta_{16} + \alpha_5 \beta_{15} + \alpha_6 \beta_{18})}{\beta_{10} \beta_{19} + \alpha_4 \beta_{13} \beta_{19}} \end{aligned}$$

y además;

$$\begin{aligned} \alpha_1 &= \frac{\beta_1 \beta_7}{\beta_0 \beta_9 - \beta_1 \beta_3} \quad ; \quad \alpha_2 = \frac{\beta_1 \beta_5 - \beta_2 \beta_9}{\beta_0 \beta_9 - \beta_1 \beta_3} \quad ; \quad \alpha_3 = \frac{-\beta_0 \beta_7}{\beta_0 \beta_9 - \beta_1 \beta_3} \\ \alpha_4 &= \frac{\beta_2 \beta_3 - \beta_0 \beta_5}{\beta_0 \beta_9 - \beta_1 \beta_3} \quad ; \quad \alpha_5 = \frac{\beta_1 \beta_6}{\beta_0 \beta_9 - \beta_1 \beta_3} \quad ; \quad \alpha_6 = \frac{-\beta_0 \beta_6}{\beta_0 \beta_9 - \beta_1 \beta_3} \end{aligned}$$

Luego aplicamos la descomposición de Jordan $\mathbf{W} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^{-1}$, donde $\mathbf{\Lambda}$ es una matriz de valores propios y \mathbf{P} una matriz de vectores propios. Usando el polinomio característico $\mu^2 - \text{tr}z(\mathbf{W})\mu + \det(\mathbf{W}) = 0$, con μ es valor propio de \mathbf{W} , se puede demostrar que:

$$\mu_1 = \frac{1}{2} \left[(\text{tr}z(\mathbf{W})) - \sqrt{(\text{tr}z(\mathbf{W}))^2 - 4 \det(\mathbf{W})} \right] < 1$$

Por otro lado, es fácil demostrar que la solución de (12) viene dado por $\widehat{A}_{t+j} = \rho^j \widehat{A}_t$. Reemplazando esta solución en (17) y haciendo uso de (16) podemos obtener la función de política para la variable de co-estado, $\widehat{\lambda}_t$. Por su parte, la función de política de \widehat{k}_t se obtiene de la primera ecuación de (15). Finalmente \widehat{c}_t y \widehat{h}_t se obtienen de (13).

Muchas veces uno está interesado en conocer las funciones de política de otras variables del modelo, por ejemplo el producto y la inversión. Estas se pueden obtener a partir de las restricciones de recursos. Entonces, al log-linealizar la función de producción $y_t = A_t k_t^\alpha h_t^{1-\alpha}$ y haciendo uso de \widehat{h}_t podemos obtener la función de política para \widehat{y}_t . De igual manera, la función de política para la inversión obtenemos a partir de log-linealizar la ecuación $i_t = \gamma_x k_{t+1} - (1 - \delta)k_t$

En resumen, de la más exógena a la más endógena, estas funciones de política se pueden escribir de la siguiente manera:

$$\begin{aligned}\widehat{A}_t &= \rho \widehat{A}_{t-1} + \varepsilon_t \\ \widehat{k}_t &= \phi_{kk} \widehat{k}_{t-1} + \phi_{kA} \widehat{A}_{t-1} + \phi_{k\varepsilon} \varepsilon_t \\ \widehat{h}_t &= \phi_{hk} \widehat{k}_{t-1} + \phi_{hA} \widehat{A}_{t-1} + \phi_{h\varepsilon} \varepsilon_t \\ \widehat{y}_t &= \phi_{yk} \widehat{k}_{t-1} + \phi_{yA} \widehat{A}_{t-1} + \phi_{y\varepsilon} \varepsilon_t \\ \widehat{i}_t &= \phi_{ik} \widehat{k}_{t-1} + \phi_{iA} \widehat{A}_{t-1} + \phi_{i\varepsilon} \varepsilon_t \\ \widehat{c}_t &= \phi_{ck} \widehat{k}_{t-1} + \phi_{cA} \widehat{A}_{t-1} + \phi_{c\varepsilon} \varepsilon_t\end{aligned}$$

Donde los coeficientes ϕ_{jk} , ϕ_{jA} y $\phi_{j\varepsilon}$, con $j = \{k, h, y, i, c\}$ dependen de los parámetros estructurales del modelo.

El sistema que representa las funciones de política podemos escribirlo en términos matriciales y en forma compacta:

$$\mathbf{x}_t = \mathbf{\Gamma}_x(\theta) \mathbf{x}_{t-1} + \mathbf{\Gamma}_\varepsilon(\theta) \varepsilon_t \quad (18)$$

Donde $\mathbf{\Gamma}_x(\theta)$ y $\mathbf{\Gamma}_\varepsilon(\theta)$ son funciones de los parámetros estructurales del modelo. Con esta ecuación, (18) se pueden obtener las respuestas de las variables endógenas ante choques en el término de error, ε_t , es este caso choques de PTF. Pero también se pueden efectuar simulaciones de ciclos si se estimado ε_t a partir de los datos.

A continuación, para la estimación bayesiana escribimos el modelo en su forma estado-espacio. La ecuación de estados lo obtenemos de la solución del modelo. Usando el hecho de que \widehat{A} y \widehat{k}_t son las variables de estado entonces podemos representarlo en \mathbf{s}_t como un vector de estado. Así, la representación estado-espacio es:

$$\mathbf{s}_t = \mathbf{A} \mathbf{s}_{t-1} + \mathbf{B} \varepsilon_t \quad (19)$$

$$\mathbf{y}_t = \mathbf{M} \mathbf{s}_t + \mathbf{u}_t \quad (20)$$

Donde $\mathbf{A} = \Phi_s(\theta)$ y $\mathbf{B} = \Phi_\varepsilon(\theta)$. La ecuación (20) se conoce como ecuación de observación o ecuación de medida que relaciona un vector de variables observables \mathbf{y}_t con el vector de estados. Además, se incorpora un error de medición denotado por \mathbf{u}_t , que se supone un ruido blanco. Así pues, dado que ε_t y \mathbf{u}_t siguen una distribución normal, entonces \mathbf{s}_t y \mathbf{y}_t también siguen una distribución normal. Ahora bien, se debe construir un pronóstico lineal de \mathbf{y}_t usando las observaciones anteriores $\mathbf{y}_{t|t-1}$, cuya diferencia $\mathbf{y}_t - \mathbf{y}_{t|t-1}$ constituye el error de pronóstico que sigue una distribución tipo normal con media cero y varianza constante Σ_t . El método que nos permite construir dicho pronóstico es el filtro de Kalman cuya salida se usa para calcular la función de verosimilitud

$$\mathcal{L}(\mathbf{y}_T|\theta) = \prod_{t=1}^T \frac{1}{\sqrt{(2\pi)^n |\Sigma_t|}} \exp \left\{ -\frac{(\mathbf{y}_t - \mathbf{y}_{t|t-1})' \Sigma_t^{-1} (\mathbf{y}_t - \mathbf{y}_{t|t-1})}{2} \right\}$$

Esta función se transforma en términos logarítmicos, de modo que:

$$\mathcal{L}(\mathbf{y}_T|\theta) = -\frac{nT}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log |\Sigma_t| - \frac{1}{2} \sum_{t=1}^T (\mathbf{y}_t - \mathbf{y}_{t|t-1})' \Sigma_t^{-1} (\mathbf{y}_t - \mathbf{y}_{t|t-1})$$

Siendo log el logaritmo natural. Entonces, el filtro de Kalman, como método recursivo, sigue los siguientes pasos:

1. Para $t = 1$ se deben establecer los valores iniciales de $\mathbf{y}_{t|t-1}$ y Σ_t .
2. Calculamos la predicción en t de $\mathbf{y}_{t|t-1}$ así como la matriz de varianzas y covarianzas Σ_t^{-1} , de la siguiente manera:

$$\begin{aligned} \mathbf{y}_{t|t-1} &= \mathbb{E}(\mathbf{y}_t | t-1) \\ \mathbf{y}_{t|t-1} &= \mathbb{E}(\mathbf{M}\mathbf{s}_t + \mathbf{u}_t | t-1) \\ \mathbf{y}_{t|t-1} &= \mathbf{M}\mathbf{s}_{t-1} \end{aligned}$$

y,

$$\begin{aligned} \Sigma_t &= \mathbb{E} [(\mathbf{y}_t - \mathbf{y}_{t|t-1})(\mathbf{y}_t - \mathbf{y}_{t|t-1})'] \\ \Sigma_t &= \mathbb{E} [(\mathbf{M}\mathbf{s}_t + \mathbf{u}_t - \mathbf{M}\mathbf{s}_{t|t-1})(\mathbf{M}\mathbf{s}_t + \mathbf{u}_t - \mathbf{M}\mathbf{s}_{t|t-1})'] \\ \Sigma_t &= \mathbb{E} \{ [\mathbf{M}(\mathbf{s}_t - \mathbf{s}_{t|t-1}) + \mathbf{u}_t][\mathbf{M}(\mathbf{s}_t - \mathbf{s}_{t|t-1}) + \mathbf{u}_t]'\} \\ \Sigma_t &= \mathbf{M}\mathbb{E}(\mathbf{s}_t - \mathbf{s}_{t|t-1})(\mathbf{s}_t - \mathbf{s}_{t|t-1})'\mathbf{M}' + \mathbb{E}(\mathbf{u}_t\mathbf{u}_t') \\ \Sigma_t &= \mathbf{M}\Sigma_t^y\mathbf{M}' + \Sigma_t^u \end{aligned}$$

3. Luego de observar los datos en t calculamos la verosimilitud:

$$\mathcal{L}(\mathbf{y}_t|\theta) = -\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma_t| - \frac{1}{2} (\mathbf{y}_t - \mathbf{y}_{t|t-1})' \Sigma_t^{-1} (\mathbf{y}_t - \mathbf{y}_{t|t-1})$$

4. Se actualizan los pronósticos de los estados en \mathbf{t} , \mathbf{s}_t , así como la matriz de varianzas y covarianzas, Σ_t^s , que vienen dados por:

$$\begin{aligned}\mathbf{s}_{t|t} &= \mathbf{s}_{t|t-1} + \Sigma_t^s \mathbf{M}^{-1} \Sigma_t^{-1} (\mathbf{y}_t - \mathbf{y}_{t|t-1}) \\ \Sigma_t^s &= \Sigma_{t|t-1}^s - \Sigma_{t|t-1}^s \mathbf{M}^{-1} \Sigma_t^{-1} \mathbf{M} \Sigma_{t|t-1}^s\end{aligned}$$

5. Ahora se calcula los pronósticos de los estados en $\mathbf{t} + 1$, $\mathbf{s}_{t+1|t}$, y su matriz de varianzas y covarianzas, $\Sigma_{t+1|t}^s$, de la siguiente manera:

$$\begin{aligned}\mathbf{s}_{t+1|t} &= \mathbb{E}(\mathbf{s}_{t+1}|\mathbf{t}) \\ \mathbf{s}_{t+1|t} &= \mathbb{E}[\mathbf{A}\mathbf{s}_t + \mathbf{B}\varepsilon_{t+1}|\mathbf{t}] \\ \mathbf{s}_{t+1|t} &= \mathbf{A}\mathbf{s}_{t|t}\end{aligned}$$

y,

$$\begin{aligned}\Sigma_{t+1|t}^s &= \mathbb{E}[(\mathbf{s}_{t+1} - \mathbf{s}_{t+1|t})(\mathbf{s}_{t+1} - \mathbf{s}_{t+1|t})'] \\ \Sigma_{t+1|t}^s &= \mathbb{E}[(\mathbf{A}\mathbf{s}_t + \mathbf{B}\varepsilon_{t+1} - \mathbf{A}\mathbf{s}_{t|t})(\mathbf{A}\mathbf{s}_t + \mathbf{B}\varepsilon_{t+1} - \mathbf{A}\mathbf{s}_{t|t})'] \\ \Sigma_{t+1|t}^s &= \mathbb{E}[\mathbf{A}(\mathbf{s}_t - \mathbf{s}_{t|t}) + \mathbf{B}\varepsilon_{t+1}] [(\mathbf{s}_t - \mathbf{s}_{t|t})' \mathbf{A}' + \varepsilon_{t+1}' \mathbf{B}'] \\ \Sigma_{t+1|t}^s &= \mathbf{A}\mathbb{E}[(\mathbf{s}_t - \mathbf{s}_{t|t})(\mathbf{s}_t - \mathbf{s}_{t|t})'] \mathbf{A}' + \mathbf{B}\mathbb{E}(\varepsilon_{t+1}\varepsilon_{t+1}') \mathbf{B}' \\ \Sigma_{t+1|t}^s &= \mathbf{A}\Sigma_t^s \mathbf{A}' + \mathbf{B}\Sigma_\varepsilon \mathbf{B}'\end{aligned}$$

6. Repetimos los pasos 2-5 para $t = 2, 3, 4, \dots, T$

Luego, se deben especificar las distribuciones previas de los parámetros, o los priors, $p(\theta)$. Esto puede hacerse tomando en cuenta un conjunto de criterios, como por ejemplo, si los parámetros son estrictamente positivos, parámetros delimitados entre cero y uno, entre otros. Las distribuciones más usadas son; distribución gamma inversa para que los parámetros sean positivos; distribución beta para parámetros delimitados entre cero y uno; distribución normal para parámetros que no están delimitados.

Habiendo establecido las distribuciones de probabilidad de los parámetros, podemos estimar la distribución posterior. Esta representa las probabilidades asignadas a diferentes valores de los parámetros después de observar los datos. Esto constituye una actualización de las probabilidades dadas por el prior, con base a la información adicional proporcionada por las variables en la muestra. El teorema de Bayes relaciona los priors y las posteriores.

$$p(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)p(\theta)}{p(\mathbf{y})}$$

Siendo $p(\theta|\mathbf{y})$ es la densidad de los parámetros condicionados a los datos (Posterior), $p(\mathbf{y}|\theta)$ es la densidad de los datos condicionados a los parámetros, $p(\theta)$ es la densidad incondicional de los parámetros (Prior) y $p(\mathbf{y})$ es la densidad marginal de los datos.

Tenga en cuenta que $p(\mathbf{y})$ no depende de los parámetros θ , de esta forma, puede tratarse como una constante, lo que nos permite escribir lo siguiente:

$$p(\theta|\mathbf{y}) \propto p(\mathbf{y}|\theta)p(\theta)$$

Definimos; $\mathcal{K}(\theta|\mathbf{y}) = p(\mathbf{y}|\theta)p(\theta)$ como el Kernel posterior. Al aplicar logaritmos:

$$\log \mathcal{K}(\theta|\mathbf{y}) = \log p(\mathbf{y}|\theta) + \log p(\theta)$$

$$\log \mathcal{K}(\theta|\mathbf{y}) = \mathcal{L}(\mathbf{y}|\theta) + \log p(\theta)$$

Ahora, suponiendo que los prior tienen distribución independiente, entonces podemos escribir.

$$\log \mathcal{K}(\theta|\mathbf{y}) = \mathcal{L}(\mathbf{y}|\theta) + \sum_{j=1}^J \log p(\theta_j) \quad (21)$$

Siendo J el número de parámetros a estimar. Esta ecuación es la que permite estimar los posteriors de los parámetros. Esto se logra usando métodos numéricos debido a la no linealidad contenidos en θ . El algoritmo Metropolis-Hastings (MH), basados en las cadenas de Markov de Monte-Carlo (MCMC), permite simular numéricamente las distribuciones posteriores. Este algoritmo sigue los siguientes pasos:

1. Se considera una distribución de salto para los parámetros del modelo:

$$J(\theta^*|\theta^{t-1}) = N(\theta^{t-1}, c\Sigma_{\theta^m})$$

2. Se determina r , la proporción del núcleo posterior evaluado en la nueva estimación propuesta sobre el núcleo posterior evaluado en la estimación prior propuesta:

$$r = \frac{\mathcal{K}(\theta^*|\mathbf{y})}{\mathcal{K}(\theta^{t-1}|\mathbf{y})}$$

3. Aceptar o rechazar θ^* como la nueva estimación de acuerdo a la siguiente regla:

$$\theta^t = \begin{cases} \theta^* & \text{con probabilidad } \min(r,1) \\ \theta^{t-1} & \text{en otro caso} \end{cases}$$