

*Modelación markoviana para identificar
la dinámica y pronóstico del índice de
producción industrial en México de 1980 a 2018*

*Markovian modeling to identify and to forecast the
dynamics of the industrial production index
from 1980 to 2018*

GUSTAVO CABRERA GONZÁLEZ¹
ADRIÁN DE LEÓN ARIAS²

- **Resumen:** En este artículo, por medio de modelación markoviana estudiamos la identificación de los estados estocásticos y pronóstico del índice mensual de producción industrial en México de 1980 a 2018. Dado que la muestra de datos está sujeta a fuertes fluctuaciones económicas y financieras, de una batería de modelos autorregresivos (lineales y con parámetros markovianos de cambio de régimen) se elige la especificación del modelo que mejor se ajusta a los datos a través del factor de Bayes. La selección del modelo provee evidencia de que las tasas de crecimiento mensual de este índice presentan parámetros (media y volatilidad) que cambian con el tiempo. Se lleva a cabo un ejercicio de pronóstico sobre el modelo markoviano de mejor ajuste a los datos. Para medir su capacidad de inferencia, se compara su eficiencia respecto de la especificación lineal autorregresiva en la misma serie de datos. Los resultados muestran que la media de los errores de pronóstico (dentro y fuera de la muestra) son menores en la especificación markoviana. La metodología bayesiana aplicada permite estimar de forma endógena e inferir de manera precisa incluso por problemas de identificación de parámetros markovianos, pequeño número de observaciones en regímenes, datos atípicos, número de regímenes e incertidumbre de parámetros sujetos a cambio de estado.
- **Palabras clave:** Índice de producción industrial, parámetros markovianos, análisis bayesiano, pronóstico.
- **Clasificación JEL:** E23; C24; C11; G17.
- **Abstract:** In this article, by Markov switching modeling we study the identification of unknown states and forecasting of the monthly industrial production index of Mexico from 1980 to 2018. Given that the data-sample is subject to strong eco-

¹ Universidad de Guadalajara, México. Autor para correspondencia. E-mail: gustavo.cgonzalez@academicos.udg.mx

² Universidad de Guadalajara, México.

conomic and financial fluctuations, from a battery of auto-regressive models (linear and Markov switching parameters), the specification that best fits to data through the Bayes factor is chosen. The model selection of the monthly growth rates index leads to parameters (mean and volatility) change over time. A forecast exercise is carried out on the Markovian model of best fit to data. To measure the accuracy on the estimation, its efficiency is compared with the linear auto-regressive models on the same data. Results provide evidence that the mean of the forecasting errors (in-sample and out-sample) are lower than those of the linear auto-regressive model. The Bayesian methodology applied allows to estimate endogenously and accurately infer, despite of identification problems of Markov switching parameters, small number of observations in regimes, atypical data, number of regimes, and uncertainty in parameters subject to switch.

- **Keywords:** Industrial production index, Markov switching, Bayesian analysis, forecasting.
- **JEL classification:** E23; C24; C11; G17.
- Recepción: 18/07/2018 Aceptación: 13/03/2019

■ *Introducción*

Para el modelado de series de tiempo en mercados emergentes, particularmente en México, se requiere tomar en cuenta que los mercados económico-financieros son muy susceptibles a fluctuaciones financieras y económicas, tanto nacionales e internacionales, así como a cambios en política monetaria, fiscal y de Gobierno. Por tanto, es insuficiente que el modelado de las series de tiempo esté basado en parámetros constantes en el tiempo, ya que los patrones empíricos observados y esperados no podrían ser bien capturados (Hamilton, 2005). Más aún, los datos en economías emergentes son muy peculiares (pequeño número de observaciones en algún régimen, cambios estructurales, datos atípicos y mercados volátiles). Por lo anterior, metodologías econométricas estándares basadas en maximización de la función de verosimilitud podrían violar algunos de sus supuestos y la posible estimación sería incierta (Frühwirth-Schnatter, 2006). El modelado bayesiano de series de tiempo con estas características de los datos es muy útil para académicos, investigadores y tomadores de decisión.

Hay una literatura amplia en modelos econométricos en series de tiempo económico-financieras que consideran parámetros sujetos a cambio de régimen con parámetros markovianos (MS) a partir de los trabajos de Hamilton (1989) y Hamilton y Susmel (1994). En las compilaciones de Diebold y Rudebusch (1999), Kim y Nelson (1999b), Hamilton y Raj (2002) y el trabajo de Frühwirth-Schnatter (2006), es posible encontrar una gran cantidad de referencias y evidencia empírica. Sin embargo, motivados con los trabajos de Frühwirth-Schnatter (2004) y Carrasco, Hu y Ploberger (2014) sobre mode-

los autorregresivos univariantes con parámetros sujetos a cambio de régimen (MS-AR); así como con Kaufmann y Frühwirth-Schnatter (2002) y Bauwens, Dufays y Rombouts (2014), sobre modelos de varianza condicional heterocedástica dependiente del estado (MS-ARCH y MS-GARCH), consideramos importante probar la existencia de los cambios de régimen, los parámetros que son sujetos a cambio y el número de estados posibles. Por ejemplo, una equivocada especificación del modelo podría conducir a diferentes implicaciones económicas considerando el modelo de mejor ajuste a los datos. En el trabajo de Frühwirth-Schnatter (2001) se discute el número de estados asumido exógenamente en el trabajo de Engel y Kim (1999) sobre del tipo de cambio real Estados Unidos/Reino Unido.

Para el caso de series económico-financieras en México utilizando modelado econométrico autorregresivo (AR), de heterocedasticidad condicional autorregresiva (ARCH y GARCH), y vectores autorregresivos (VAR) hay una gran cantidad de trabajos. Por tanto, solo mencionamos algunos estudios recientes que incluyen parámetros de cambio de régimen que responden diversas cuestiones económico-financieras y que difieren de aquellos resultados utilizando modelos con parámetros constantes: producción y ciclos económicos (Mejía-Reyes & Díaz-Carreño, 2010; Mejía-Reyes, Martínez Gómez & Rendón Balboa, 2005); cuestiones en series financieras (López-Herrera, 2004; López-Herrera, Ortiz-Arango y Venegas-Martínez, 2011; López-Herrera & Venegas-Martínez, 2011; Heath & Kopchak, 2015); tópicos en tipo de cambio (Bazdresch & Werner, 2005; Benavides & Capistrán, 2012; Islas-Camargo, Cortez y Sanabria Flores, 2017). Sin embargo, en todos estos trabajos la apropiada especificación del modelo no es discutida formalmente. Una excepción es el estudio de Cabrera, Coronado, Rojas y Romero-Meza (2018), donde se analizan los cambios de volatilidad en el índice de precios y cotizaciones de México, encontrando tres estados por el factor de Bayes. No obstante, este artículo omite cualquier ejercicio de pronóstico dentro y fuera de la muestra del índice de precios y de la volatilidad asociada.

Es importante mencionar que la especificación sugerida en nuestro trabajo no necesariamente difiere de las impuestas exógenamente en los trabajos mencionados anteriormente ni tampoco contradice sus hallazgos. El objetivo del presente trabajo es complementar y posiblemente coincidir con algunas de las especificaciones asumidas, pero a diferencia de estos, el modelo econométrico sugerido en el presente trabajo es endógeno más que impuesto a priori por el investigador.

Por otro lado, una cuestión siempre importante para el modelado econométrico es que se considere el modelo de mejor ajuste de los datos y no solo mostrar que una especificación diferente deja resultados e implicaciones económicas distintas, sino también es igual de relevante el pronóstico de las series dentro y fuera de la muestra principalmente en modelos de cambio de régimen MS (Engel, 1994). Para esto, una vez que se elige una especificación sugerida por el factor de Bayes con base en Kass y Raftery (1995) y Frühwirth-Schnatter (2004), se realiza un ejercicio de pronóstico fuera de la muestra sobre las especificaciones MS, donde los errores de pronóstico se comparan con enfoques alternativos incluyendo los modelos lineales cuyos errores son muy difíciles de mejorar en la mayoría de las aplicaciones empíricas. Un tema reconocido como

una debilidad de las especificaciones MS (Boot y Pick, 2017). Nuestro trabajo es muy cercano en metodología econométrica al de Cabrera, Coronado, Rojas y Romero-Meza (2018), pero aquí exploramos la pregunta sobre si el modelo de mejor ajuste MS deja errores de pronóstico fuera de la muestra menores que su principal competidor lineal.

En este contexto, la contribución de este trabajo consiste en determinar si la serie de tiempo del índice mensual de producción industrial (IP) en México, no solo tiene comportamientos no lineales, sino cuál sería el modelo econométrico apropiado de acuerdo con su mejor ajuste a los datos para analizar su dinámica e inferencia. Una vez conocida esta especificación, mostrar si se pueden estimar errores de pronóstico menores que su principal competidor, el modelo lineal AR.

Finalmente, consideramos importante mencionar que este artículo no es una simple batería de pruebas de no linealidad que se aplican a series económicas y dan una posible explicación del resultado (por ejemplo, los errores estocásticos estimados y sus cuadrados no son independiente e idénticamente distribuidos). Estas pruebas solo deberían ser una premisa para motivar el modelo econométrico no lineal. Tampoco se tiene como objetivo mostrar que algunas series de tiempo han sido mal especificadas y por tanto las conclusiones derivadas de esto podrían ser diferentes. Nuestro trabajo intenta complementar los hallazgos anteriores. Por último, también se desea dejar evidencia de que pruebas de cointegración, raíces unitarias, modelos con parámetros constantes VAR, ARCH y GARCH como supuestos de normalidad e independencia en los errores al cuadrado, deberían considerarse con precaución en las series de tiempo similares a la estudiada en este artículo.

En particular, asumir una especificación MS a priori, el caso estándar de dos estados como la mayoría de las referencias mencionadas lo asume, tendrá implicaciones económicas diferentes si solo hay un estado o si hay tres o más. Por ejemplo, un tercer estado podría capturar fuertes fluctuaciones o datos atípicos que se presentan en menor porcentaje de la muestra. En dichos periodos de tiempo, la volatilidad es alta y muy probablemente sin dinámica autorregresiva. Lo que podría dejar en términos financieros un mercado eficiente (no es posible predecir el precio de un activo con información del pasado). La volatilidad alta también implica riesgo e inestabilidad financiera quizás en breves periodos de tiempo, lo cual es información importante para los tomadores de decisión e inversionistas. Por otro lado, un modelo mal especificado de dos estados capturaría dichos datos en un solo estado con mayor persistencia, cuando el segundo régimen pueda solo corresponder a volatilidad media en donde fuera posible encontrar periodos no eficientes en el mercado, pero al incluir estos datos del tercer estado en el segundo, podría cambiar la eficiencia del tercero.³

Si bien es posible mostrar que un modelo de dos estados es siempre posible de estimarse bajo restricciones en los parámetros y valores iniciales cercanos a un máximo local, es de observarse que omitiendo dichas restricciones podría modificar el máximo y no existir cambios de estado o quizá más de dos. Las implicaciones económicas

³ Véase Islas-Camargo, Cortez y Sanabria Flores (2017) para una discusión de eficiencia de dos regímenes en el mercado del tipo de cambio en México.

derivadas de una errónea especificación del modelo estarán condicionadas a la especificación impuesta por el investigador. Por lo tanto, consideramos una herramienta de selección de especificación de mejor ajuste de datos útil para tomadores de decisiones, análisis económico e inferencia dentro y fuera de la muestra. Por otro lado, información sobre los parámetros sujetos a cambio de régimen proveen evidencia importante de riesgo en precios de los activos. Por ejemplo, una media de cero con volatilidad grande podría ser equivalente a una media alta con volatilidad baja y una media negativa con volatilidad baja. En el primero de los casos, el tomador de decisiones tendría sus precauciones ante volatilidad alta, en los últimos el riesgo sería menor. Sin embargo, si ambos parámetros son sujetos a cambio de régimen es posible determinar si hay media y/o volatilidad cambiante en el tiempo, lo cual determinaría de forma más precisa la decisión del inversionista.

Finalmente, es importante no omitir que este trabajo no busca determinantes exógenos de la dinámica del IP. Esto se deja para futura investigación, considerando la premisa de que un posible modelo estructural debería considerar parámetros MS. El ejercicio de pronóstico fuera de la muestra es basado en el enfoque de pesos óptimos en modelos MS de Boot y Pick (2017); esto no implica que el pronóstico de la serie o sus errores no puedan ser mejorados incluyendo determinantes exógenos en el modelo, o quizá como en las series financieras, el proceso sea eficiente y no haya forma de predecir su rendimiento en el futuro, el único objetivo es mostrar si es posible mejorar los errores de pronóstico en comparación con su principal competidor lineal.

Respecto de la cuestión metodológica, es relevante mencionar que de acuerdo con Frühwirth-Schnatter (2006), los métodos de simulación bayesiana por cadenas de Markov Monte Carlo (MCMC), en el caso particular de series de las series de tiempo cortas y con posiblemente pocos datos en algún régimen, dejan una estimación más precisa que utilizar el enfoque frecuentista de la maximización de la verosimilitud (MLE). Mas aún, bajo una adecuada selección de las distribuciones de probabilidad a priori, la MLE es un caso particular del MCMC. En la sección de metodología discutimos por qué preferimos el enfoque bayesiano para este trabajo.

Para presentar los resultados de nuestra investigación, el artículo se organizará de la siguiente manera: en la segunda sección se discutirán algunas propiedades estadísticas de los datos; en la tercera sección la estructura econométrica y la metodología bayesiana serán descritas y adicionalmente, se justificara la aplicación de esta metodología; en la cuarta sección se discutirán los resultados y se llevara a cabo el ejercicio de pronóstico del índice de producción industrial en México. Finalmente, se presentarán las conclusiones en la última sección.

■ *Datos del índice de producción en México*

Esta sección ofrecemos una perspectiva exploratoria de los datos que sustentan la modelación econométrica de la siguiente sección. Para esto, se analizan de manera gráfica y en un análisis exploratorio los datos mensuales del IP de México del mes enero de 1980 al a marzo de 2018. Se ha elegido el largo de la serie (460 meses) de acuerdo

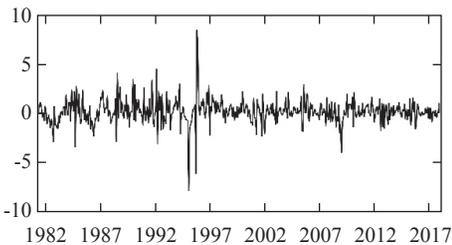
con la disponibilidad de los datos incluyendo los periodos de recesiones económico-financieras más importantes de México.⁴

En la Figura 1 graficamos las tasas de crecimiento porcentual del IP junto con la estimación de su densidad kernel de probabilidad subyacente.⁵ De la función de ajuste por el kernel, no es claro una mezcla de distribuciones normales, más bien, la frecuencia mensual de los datos deja una distribución con media cercana a cero, pero con largas colas. Sin embargo, parece tener cambios en la volatilidad de la serie. La representación de la serie de las tasas de crecimiento del IP podría sugerir, volatilidad alta entre 2008-2009, media antes de 2008 y baja después de 2009. Para dicha serie se modelarán hasta la posible existencia de cuatro estados.

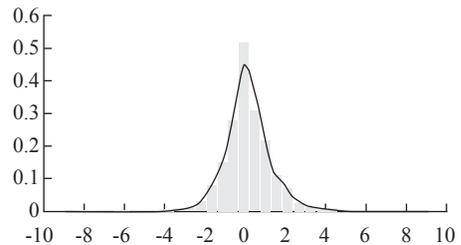
Un punto fino para denotar de este análisis exploratorio es el número de datos en cada posible estado. Es muy probable que en las recesiones profundas solo haya tres o cuatro datos por cada una de ellas. Esto viola los supuestos de MLE y muy probablemente la estimación podría fallar.⁶

De la Figura 1A la serie de tiempo en tasas porcentuales es claramente estacionaria. De los momentos muestrales, la inspección gráfica muestra que la varianza no es constante en el tiempo. La serie del IP como todos los índices económicos tiene tendencia y no tiene momentos constantes en el tiempo. De la Figura 1B la prueba de normalidad se rechazaría por las colas largas de la distribución.⁷

Figura 1
Gráfica descriptiva del IP en tasas de crecimiento mensual



A) Tasa de crecimiento mensual del índice de producción industrial en porcentaje.



B) Histograma y función de densidad con base en el kernel de una normal.

Fuente: Elaboración propia.

⁴ La fuente de los datos es la Organización para la Cooperación y el Desarrollo Económicos (OECD), (<http://oecd.org>).

⁵ La densidad kernel se estimó con Matlab a través de la función `ksdensity`. Dicha función está basada en el *kernel* de una distribución normal.

⁶ Es importante notar que la estimación del modelo de cambio de régimen Markovianos por MLE en STATA o cualquier otro paquete econométrico para esta serie del IP (en el caso de tres estados o más), tiene problemas porque el algoritmo de optimización no converge.

⁷ Dadas las limitaciones del espacio no se presenta la estadística descriptiva, pero se encuentra disponible a petición en la dirección electrónica de los autores.

■ *Modelado econométrico y método de estimación bayesiano*

En la primera parte de esta sección, describimos el modelo econométrico para estudiar las tasas de crecimiento mensual del IP en México. El modelo propuesto es un autorregresivo con MS, tanto en la media como en la varianza, y cuyo caso particular es el modelo lineal con parámetros constantes.

La notación y definiciones están basadas en Timmermann (2000), Frühwirth-Schnatter (2004), y aplicada antes en Cabrera, Coronado, Rojas y Romero-Meza (2018). Sea y_t la serie de tiempo estacionaria en tasas de crecimiento de datos mensuales. El modelo base es un autorregresivo en donde todos sus parámetros siguen cambios de régimen (denotado como MSAR) dado por:

$$(1) \quad \phi_{[s_t]}(L)y_t = \zeta_{[s_t]} + \sigma_{[s_t]}\varepsilon_t$$

donde $\varepsilon_t \sim iidN(0, 1)$; los parámetros autorregresivos $\phi_{[s_t]}$ dependen del estado, el orden del polinomio es el número de rezagos (p) y L es el operador de rezagos; el intercepto $\zeta_{[s_t]}$ implicaría una media dependiente del estado, y $\sigma_{[s_t]}$ captura la volatilidad del término de error por estado. El indicador de los estados S_t es una variable estocástica definida en un espacio discreto y finito $\{1, \dots, K\}$ que captura la dependencia de los parámetros del régimen para todo t . En dicho indicador, se asume que sigue un proceso de Markov de primer orden; es decir, la probabilidad de transición del estado k al estado l esta dado por $\xi_{kl} = \Pr(S_t = l | S_{t-1} = k)$ para $t=1, \dots, T$, y los estados $k, l \in \{1, \dots, K\}$. El modelo autorregresivo (1) tiene como casos particulares los siguientes:

- Si solo el intercepto $\zeta_{[s_t]}$ y la volatilidad del término de error $\sigma_{[s_t]}$ dependen del estado, el modelo (1) se reduce al modelo denotado como MSIAR:

$$(2) \quad \phi(L)y_t = \zeta_{[s_t]} + \sigma_{[s_t]}\varepsilon_t$$

- Si solo hay un estado, $K=1$, el modelo (1) se reduce al modelo lineal autorregresivo denotado como LAR:

$$(3) \quad \phi(L)y_t = \zeta + \sigma\varepsilon_t$$

En resumen, para las tasas de crecimiento mensual del índice de producción, de acuerdo con todas las posibles combinaciones de parámetros y estados se tienen $K(2p+1)-p$ posibles especificaciones autoregresivas definidas como MSAR, MSIAR y LAR, una vez estimados entonces se procede a elegir el modelo de acuerdo con su ajuste de datos. El método de simulación bayesiana de Frühwirth-Schnatter (2004), nos permitirá llevar a cabo este objetivo. En particular, la ecuación (1) arroja una media y una volatilidad que dependen de los parámetros autorregresivos y el intercepto en cada estado. En el (2) los momentos dependen del intercepto por estado y en el último, (3)

los momentos son constantes en el tiempo. Los momentos de cada uno de estos modelos son estudiados en Timmermann (2000).

Factor de Bayes y la probabilidad marginal (ML)

Usando la técnica del muestreo puente (*Bridge-sampling*) de Meng y Wong (1996) y los métodos de simulación de MCMC con base en Frühwirth-Schnatter (2004, 2006), es posible estimar de manera precisa la probabilidad o verosimilitud marginal (ML); es decir, $\Pr(y|M_j)$, y a partir de esta estimar la probabilidad de cada modelo condicional a la muestra de datos, $\Pr(M_j|y)$. De acuerdo con la literatura, el factor de Bayes para comparar el modelo i con el modelo j esta dado por:

$$BF_{ij} = \frac{\Pr(y|M_j)}{\Pr(y|M_i)} = \frac{\Pr(M_j|y)}{\Pr(M_i|y)}$$

donde o hay ninguna probabilidad preferida a priori para algún modelo; esto es, $\Pr(M_j) = \Pr(M_i)$. Por lo tanto, estimando la ML de cada especificación, el modelo con mayor tasa podría ser preferido de acuerdo con la guía sugerida en Kass y Raftery (1995).

En la Tabla 1, que ofrece la guía convencional para la selección del modelo con base en el factor de Bayes, las probabilidades de los datos condicionales al modelo i respecto del modelo j con más alta ML son calculados para diferentes valores del logaritmo del factor de Bayes. Por ejemplo, si el $\log_e(BF_{ij}) = 7$ o mayor de 3 en base 10, el modelo j tiene un soporte de 99.9% de los datos; por lo tanto, se considera evidencia positiva en contra del modelo i y se puede rechazar. Por otro lado, si el logaritmo base 10 del factor de Bayes está entre -3 y 3 no es posible descartar ninguna especificación y ambos modelos pueden representar a los datos.

Es importante mencionar que en el artículo de Kass y Raftery (1995) se muestra que los criterios Akaike y Schwarz para elegir el número de rezagos en procesos lineales autorregresivos son asintóticamente equivalentes al factor de Bayes. Más aún, los autores presentan una correspondencia con la tabla chi-cuadrada de estadística clásica para referencia en pruebas de razón de verosimilitud utilizadas en econometría frecuentista para elegir la especificación en modelos anidados.

Procedimiento de estimación bayesiana

En este apartado describimos de manera general el procedimiento de estimación por métodos de simulación bayesiana para inferir de manera precisa la selección del modelo con mejor ajuste a los datos y llevar acabo las inferencias asociadas del modelo seleccionado.

La simulación de la ML de cada modelo autorregresivo se lleva a cabo con base en Frühwirth-Schnatter (2004), los detalles y el código de MATLAB se pueden encontrar en el trabajo de Frühwirth-Schnatter (2006). Por lo tanto, el factor de Bayes para cada combinación posible de modelos puede ser estimado dada la razón en logaritmos de las ML.

Tabla 1
Guía de selección del factor de Bayes

$\log_e(BF_{ij})$	$\Pr(M_i y)$	$\Pr(M_j y)$	$\log_{10}(BF_{ij})$	Guía de selección
7	0.999	0.001	3.040	Evidencia positiva en contra del modelo i
3	0.953	0.047	1.303	
1	0.731	0.269	0.434	
0	0.500	0.500	0.000	Nada que decir ambos modelos pueden representar los datos
-1	0.269	0.731	-0.434	
-3	0.047	0.953	-1.303	
-7	0.001	0.999	-3.040	Evidencia positiva en contra del modelo j

Nota: Si el factor de Bayes en logaritmo base 10 es mayor de 20 (en valor absoluto) la evidencia es fuerte.

Fuente: Elaboración propia con base en Kass y Raftery (1995).

Una vez que elegida la especificación, se lleva a cabo un nuevo MCMC restringido. Esto es, en el método de simulación Monte Carlo se imponen las restricciones de identificación para los parámetros dependientes del estado. Es decir, si la especificación autorregresiva preferida es un modelo de tres estados, una posible restricción para el muestreo MCMC es: $\sigma_1^2 > \sigma_2^2 > \sigma_3^2$ y $\mu_1 > \mu_2 > \mu_3$. El procedimiento del MCMC restringido es como sigue:

- Para valores iniciales de los indicadores de los estados S y varianzas para el término de error, muestrear los parámetros ψ condicional a los datos y S , desde la función $p(\psi|y, S)$. Los parámetros autorregresivos se muestrean a partir de un normal multivariante, la varianza de una gamma invertida, y las probabilidades de transición de una función de densidad de probabilidad Dirichlet.
- Posteriormente, aplicar el muestreo rechazo-aceptación al paso anterior imponiendo las restricciones de identificación dependientes del estado. Esto es, si los parámetros muestreados satisfacen las restricciones guarda los parámetros muestreados, de lo contrario se mantienen los anteriores.
- Dados los parámetros del último paso, muestrear los indicadores de los estados por el algoritmo *forward-smoothed-backward* sugerido por Kim y Nelson (1999a) y descrito en Frühwirth-Schnatter (2006).
- Repetir todos los pasos anteriores 7 000 veces descartando los primeros 2 000 para eliminar la dependencia de los valores iniciales.

Dada la salida del MCMC restringida, se pueden calcular todos los momentos posteriores para todos los parámetros (por ejemplo, la media y la desviación estándar). Así mismo, de la salida del MCMC restringida y de los indicadores simulados por el algoritmo *forward-smoothed-backward*, se estima la media posterior de las probabilidades suavizadas de cada estado para $t = 1, \dots, T$ (en la sección de resultados, éstas serán graficadas), lo cual es una de las contribuciones más importantes de los modelos con cambio de régimen.

¿Por qué aplicar métodos bayesianos a diferencia de estimación MLE?

La respuesta está basada en dos cuestiones: eficiencia en la estimación y permite la comparación entre especificaciones markovianas. En primer lugar, es importante recordar por qué los métodos bayesianos podrían superar la eficiencia en la estimación respecto de la econometría frecuentista por maximización de la verosimilitud (MLE) por las siguientes razones: primero, en la estimación por MCMC las propiedades asintóticas aplican al número de simulaciones más que al tamaño de la muestra; segundo, no es necesario las condiciones de segundo orden ni la inversa del valor esperado de la matriz de información, que en el caso de MLE, es muy común que dicha matriz sea singular en las iteraciones del algoritmo de maximización. Los segundos momentos se derivan de la salida del MCMC sin optimizar. Tercero, para encontrar la convergencia en el máximo local y/o global en MLE se requieren muy buenos valores iniciales y acotar el dominio de maximización imponiendo restricciones, el caso del MCMC no depende de los valores iniciales para la convergencia de la cadena sino más bien de las distribuciones de probabilidad a priori que aseguren una forma conocida de la distribución de probabilidad condicional posterior. Así también, un algoritmo que permita visitar el dominio del parámetro en la mayor medida posible. Por lo anterior, la estimación bayesiana podría dejar una estimación más precisa que MLE si la optimización falla por alguna de las razones mencionadas anteriormente. Mas aún, bajo una adecuada elección de las distribuciones a priori, la estimación MLE es un caso particular del MCMC.

En segundo lugar, describiremos los motivos por los cuáles no es posible comparar modelos markovianos utilizando los criterios estándar de la econometría frecuentista a diferencia del enfoque bayesiano. La estimación lineal autorregresiva y comparación de modelos (es decir, el número de rezagos), es casi siempre factible aplicarla por la razón de verosimilitudes de cada modelo. Los criterios de la econometría frecuentista son Akaike, Schwarz y Hannan-Quinn. Todos utilizan el estadístico de contraste que sigue una distribución chi-cuadrada y en el caso de muestras de datos relativamente grandes, las propiedades asintóticas dejan estimadores insesgados y la prueba no tiene ningún problema.

Por otro lado, es importante notar que, para probar un modelo con dos estados en contra de un modelo con un solo estado (lineal) la razón de las verosimilitudes no se aproxima a una chi-cuadrada ni asintóticamente. La razón es por dos cuestiones: primero, si bajo la hipótesis nula el verdadero modelo es lineal, cualquier otro modelo con parámetros MS no está identificado; este problema es llamado en la literatura *nuisance-parameters*,⁸ Es decir, si solo hay una varianza o una media, no hay una forma única de estimar dos o más valores de los parámetros. Una técnica muy frecuentemente utilizada es aplicar restricciones a la optimización; por ejemplo, en el caso de dos estados: una media es más grande que la otra y/o una varianza es mayor que la otra. La estimación es factible y MLE encontrará un máximo local. Pero si cambiamos las restricciones es probable que MLE encuentre otro máximo local. Por ejemplo: si se estima el modelo sin restricciones, una posibilidad es que una media sea diferente de cero y la otra no, un modelo equivalente sería la primera es cero y la segunda no, pero otro modelo equi-

⁸ Véase Andrews y Ploberger (1994) y Hansen (1996).

valente es, alguna suma ponderada de las medias es igual a la media del modelo lineal. En el caso de dos varianzas esto no puede ocurrir porque una varianza no puede ser cero en ningún estado, pero si es posible que alguna media ponderada de dos varianzas equivalga a la varianza del modelo lineal.

Segundo, los errores estándar pueden obtenerse con la inversa del valor esperado de la negativa de la matriz de información (matriz de segundas derivadas de la función de verosimilitud); sin embargo, en el caso de que la hipótesis nula no sea rechazada esta matriz suele ser singular. Mas aún, la matriz tiende a ser singular también en el caso de pocas observaciones en un estado.⁹

El artículo de Carrasco, Hu y Ploberger (2014) resuelve el problema antes mencionado de manera parcial. Esto es, si únicamente se estima el modelo lineal es posible aplicar una batería de pruebas estadísticas para comprobar si el parámetro es dependiente del estado. La prueba es asintótica y puede realizarse por MLE al modelo lineal. Sin embargo, no puede efectuarse para comparar dos estados contra tres estados, o un modelo en donde solo la media cambia contra otro en donde la media y la varianza son dependientes del estado. Lo anterior, porque las pruebas sugeridas en el artículo tienen como base el modelo lineal.

Por otra parte, los enfoques bayesianos utilizados en Frühwirth-Schnatter (2004) y Bauwens, Dufays y Rombouts (2014) permiten la comparación de los modelos con cualquier número de estados, aunque los parámetros no estén identificados ni tomando como base única el modelo lineal. Lo anterior se debe a dos razones: primero, los errores estándar y todos los momentos de los parámetros pueden estimarse por la salida del MCMC; es decir, no hay optimización. En segundo lugar, si un parámetro no está identificado los muestreos del parámetro coincidirán en dos estados y entonces se evaluará la función de probabilidad de dos parámetros contra uno. Esto dejará que la función de probabilidad univariante sea más alta que una bivalente y con esto se concluiría que no hay dos parámetros. Lo mismo ocurriría con cualquier número de estados. En la literatura es común restringir el muestreo aplicando aceptación-rechazo. Es decir, guardar el muestreo si una media es más grande que la otra, en otro caso, descartar el muestro. Sin embargo, al final del muestreo es posible observar que las salidas del MCMC estarán truncadas y no será posible estimar la función de probabilidad para todo el dominio, es decir, solo se evaluaría para la parte de muestreos aceptados. Esto dejará que la estimación de la probabilidad del parámetro sea sesgada, por lo tanto, la probabilidad de todo el modelo (ML) será errónea bajo este problema. No obstante, si se impone la restricción y los muestreos cubren una gran parte del dominio del parámetro (es decir, sin truncamiento) entonces si habría dos parámetros claramente separados y el problema de sesgo disminuye. El problema de truncar la función es importante cuando un parámetro es cercano a otro en dos estados; es decir, cuando quizás el modelo es lineal. En resumen, ambos trabajos confirman que, bajo una estimación sin imponer restricciones de identificación al muestreo, las ML de cada modelo siempre podrán ser comparadas. A partir de estas entonces se puede aplicar el criterio de selección de la mejor especificación de ajuste a los datos por el factor de Bayes.

⁹ Véase Hamilton (1994) y Frühwirth-Schnatter (2006).

■ Resultados

Esta sección la conformamos de tres apartados: primero, presentamos los resultados de la selección en la especificación del modelo de mejor ajuste a los datos con base en el factor de Bayes; segundo, reportamos las inferencias asociadas al modelo elegido; finalmente presentamos en un apartado los resultados del ejercicio de pronóstico dentro y fuera de la muestra del modelo seleccionado en contraste con sus principales competidores.

La Tabla 2 muestra los resultados de aplicar el factor de Bayes a las tasas de crecimiento mensual del IP en México. Con base en la simulación de la ML descrita en la sección anterior, se toma como referencia el modelo que tiene la más alta ML y contra este se calcula el factor de Bayes de cada modelo. En la Tabla 2 se resalta con números en negritas el modelo con factor de cero puesto que se compara con el mismo; es decir, MSAR de cuatro estados y un solo rezago. De acuerdo con la Tabla 1 de referencia del factor de Bayes, los modelos sugeridos para el IP son todos aquellos de tres estados bajo especificaciones MSAR y MSIAR, y con cuatro estados y uno y dos rezagos marcados sus números con negritas. El resto de los modelos son rechazados con evidencia positiva (mayor de 3) y evidencia fuerte (mayor de 20) en el caso de especificaciones lineales.

La estimación de la probabilidad marginal por simulación MCMC está basada en Frühwirth-Schnatter (2006). La estimación del factor de Bayes (en logaritmo base 10) para el IP en tasas de crecimiento mensual, es con base en el modelo con probabilidad marginal más alta. Las especificaciones marcadas en negritas son modelos sugeridos por el factor de Bayes con base en la Tabla 1. Es relevante notar que series de tiempo económicas con cuatro estados no son nuevas en la literatura, en el trabajo de Frühwirth-Schnatter (2001) sobre el tipo de cambio real Estados Unidos/Reino Unido se encontraron cuatro en lugar de los tres estados asumidos exógenamente en Engel y Kim (1999). Cada uno de estos regímenes se vinculó con diversas causas económico-financieras en estos países. Por otra parte, Clements y Krolzig (1998) encuentran tres estados para PIB de Estados Unidos.

Tabla 2
Modelo y el factor de Bayes para el IP industrial de México

Rezagos/Estados	MSAR				MSIAR		
	1	2	3	4	2	3	4
0	30.62	5.30	2.33	3.27			
1	31.39	6.41	2.14	0.00	5.93	2.66	1.40
2	31.37	4.97	1.98	2.77	4.35	1.33	2.21
3	32.03	5.70	2.61	4.12	4.93	2.06	3.67
4	32.50	5.50	2.84	5.62	4.57	1.88	3.95

Nota: La razón del factor de Bayes está medida en logaritmo base 10 en relación con el modelo de más alta probabilidad marginal, es decir, la especificación MSAR (1).

Fuente: Elaboración propia.

En las tasas de crecimiento mensual del IP, de acuerdo con su gráfica, se esperaba encontrar tres estados para diferentes volatilidades; sin embargo, el factor de Bayes sugiere también cuatro estados con un rezago para el modelo MSAR definido en la ecuación (1) con mayor probabilidad marginal. Con el fin de estudiar las propiedades del modelo de cuatro estados, se elige dicha especificación con un solo rezago. La estimación del modelo se lleva a cabo por el MCMC restringido descrito en la sección anterior.

En la Tabla 3 se reportan los momentos posteriores; es decir, la media y el intervalo de alta densidad posterior a 95% denominado como HPDI¹⁰ de cada uno de los parámetros del modelo. De la media posterior de los parámetros de cada estado, se pueden derivar la media y la volatilidad del proceso. Los interceptos de cada estado son diferentes de cero excepto para el estado dos, ya que el HPDI incluye el cero (-0.3187 y 0.2719). Todos los autorregresivos son diferentes de cero excepto el caso del estado uno de acuerdo con el HPDI (-0.4723 y 0.7938). Los momentos de las tasas de crecimiento del IP son: en el primer estado, media negativa alta (-4.4355) con volatilidad de 2.3796; en el segundo, media de cero con volatilidad 1.1497; en el tercero la media es de 0.1784 con volatilidad más baja 0.9695, y en el último estado la media y la volatilidad son altas, 0.8853 y 1.8327. Las volatilidades se calcularon por las ecuaciones Yule-Walker condicionales a los estados.

Por otro lado, por la estimación de las probabilidades de persistencia de la matriz de transición¹¹ podemos concluir que la más alta es la del tercer estado; es decir, permanecer en este régimen de crecimiento moderado con la volatilidad más baja es 0.9571. La probabilidad menos persistente se refiere a volatilidad alta y recesión profunda 0.4727, lo cual coincide con las recesiones económico-financieras como se verá más adelante en el gráfico de probabilidades suavizadas en la Figura 2. Finalmente, de las probabilidades de estado estacionario se puede concluir que 2% del tiempo hay recesión profunda, 23% sin crecimiento, 66% en crecimiento moderado y 9% crecimiento alto aproximadamente (véase el área gris de la Figura 2).

En la Figura 2 se grafica las probabilidades suavizadas para todo t de cada estado junto con las tasas de crecimiento del IP. En la Figura 2A el estado uno se caracteriza por alta volatilidad con y profunda recesión decrecimiento, el tiempo de las probabilidades más altas coincide con las recesiones económicas de la década de 1980, 1995 y 2008-2009. En el segundo estado (Figura 2B) el crecimiento es nulo y la volatilidad es media, los periodos de tiempo coinciden con etapas de los ochenta y periodos inmediatos después de las recesiones de 1995 y 2008-2009. El tercer estado (Figura 2C) es el más persistente y el que se presenta en mayor porcentaje en los datos, con tasa de crecimiento moderada y la volatilidad más baja. Finalmente, en la Figura 2D el cuarto estado captura apropiadamente los crecimientos fuertes después de las crisis económicas de los ochenta y en periodos de tiempo de los noventa.

¹⁰ El HPDI (por sus siglas en inglés), se presenta el muestreo en el lugar 2.5 y 97.5% de cada parámetro de la salida del MCMC, el cual es el equivalente al intervalo de confianza de 95% en econometría frecuentista.

¹¹ Dadas las limitaciones de espacio también se omitieron el resto de las probabilidades de transición, no obstante están disponibles a petición en la dirección electrónica de los autores.

Tabla 3
Momentos posteriores del modelo MSAR ecuación (1)

Parámetro	Media	HDPI al 95%		Parámetro	Media	HDPI al 95%	
$\zeta_{S_t=1}$	-4.1254	-5.3150	-2.7709	ρ_{11}	0.4727	0.0556	0.9558
$\phi_{1,S_t=1}$	0.0699	-0.4723	0.7938	ρ_{22}	0.8588	0.1752	0.9817
$\sigma_{S_t=1}^2$	5.6346	1.3881	17.8508	ρ_{33}	0.9571	0.7425	0.9996
$\zeta_{S_t=2}$	-0.0189	-0.3187	0.2719	ρ_{44}	0.6718	0.0636	0.9697
$\phi_{1,S_t=2}$	0.4085	0.2488	0.6250	Probabilidades de estado estacionario			
$\sigma_{S_t=2}^2$	1.1012	0.8306	1.5428	ρ_1	0.0227	ρ_3	0.6539
$\zeta_{S_t=3}$	0.2273	0.0346	0.4170	ρ_2	0.2333	ρ_4	0.0901
$\phi_{1,S_t=3}$	-0.2740	-0.4311	-0.0690	Media y volatilidad de cada estado			
$\sigma_{S_t=3}^2$	0.8694	0.6746	1.0719	μ_1	-4.4355	σ_1	2.3796
$\zeta_{S_t=4}$	1.5328	0.2199	2.5684	μ_2	-0.0319	σ_2	1.1497
$\phi_{1,S_t=4}$	-0.7314	-0.9692	-0.5117	μ_3	0.1784	σ_3	0.9695
$\sigma_{S_t=4}^2$	1.5621	0.6661	3.3786	μ_4	0.8853	σ_4	1.8327

Nota: Los momentos y el HDPI posteriores se obtienen de la salida del MCMC restringido descrito en la sección 3.2. El subíndice indica el estado del parámetro de la media y la volatilidad.
Fuente: Elaboración propia.

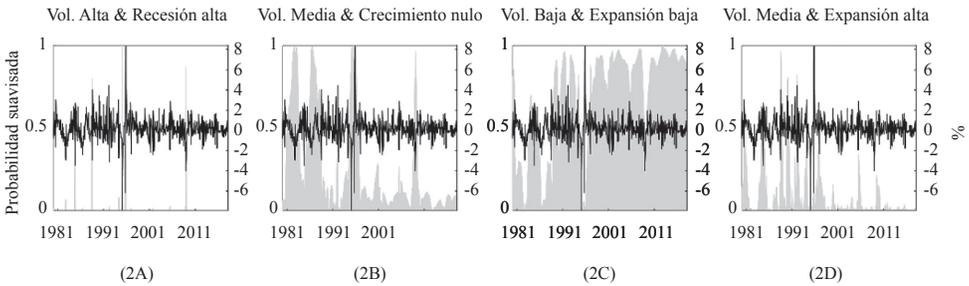
De estos gráficos de probabilidades suavizadas dos aspectos importantes a denotar son los siguientes: primero, el estado de expansión alta no se vuelve a presentar con una probabilidad importante de 2000 hasta 2018; segundo, a partir de 1995, parece darse un cambio estructural y solo hay probabilidad importante de dos estados con alta probabilidad: volatilidad media y crecimiento nulo y volatilidad baja y expansión baja. Es decir, se puede asumir a partir de 1995 un solo un estado y podría considerarse una variable dicotómica para 2008-2009. Tercero, aunque este último resultado en los cambios de estado podría considerarse esperado, los cambios son inferidos endógenamente y tienen un sustento estadístico. De los cuáles hay una probabilidad de volverse a presentar en algún momento del tiempo, principalmente los estados dos y tres.

Pronóstico en modelos de cambio de régimen MS

En esta sección concluimos con un ejercicio de pronóstico a partir del modelo seleccionado por el factor de Bayes con más alta ML para el IP. Lo anterior, con el fin de mostrar si el error de pronóstico por esta especificación MS de mejor ajuste a los datos, tiene alguna ventaja sobre su principal competidor, el modelo lineal autorregresivo. Sin embargo, dejamos para trabajos posteriores el pronóstico de la volatilidad con base en modelos MS-ARCH y MS-GARCH.

El procedimiento estándar de pronóstico en modelos MS, implica utilizar las probabilidades suavizadas de la Figura 2 y de transición como estacionarias para tomar una media ponderada y dar un pronóstico de la serie (este se denotará como MS). Sin embargo, en la mayoría de los casos los errores de pronóstico de modelos MS no pueden superar a los del modelo lineal autorregresivo. Lo anterior, ya que los momentos

Figura 2
 Probabilidades suavizadas y el índice de producción industrial (porcentaje)



Nota: Las probabilidades suavizadas fueron estimadas de la media posterior de la salida del MCMC restringido.
 Fuente: Elaboración propia.

no condicionales de modelos MS equivalen a los momentos del lineal autorregresivo. Este dilema y las referencias son discutidas en Boot y Pick (2017). En dicho artículo, se sugiere utilizar óptimamente pesos para toda la serie con base en una especificación MS y aplicarlos para pronosticar la serie en lugar de las probabilidades suavizadas. Los autores concluyen que los pesos óptimos son elegidos para minimizar el error cuadrático de pronóstico (MSFE) y al aplicarlos en toda la serie, dejan errores de pronóstico dentro y fuera de la muestra menores que los del modelo lineal. Esto solamente en el caso de que la serie de tiempo presente cambios de régimen como proceso generador de datos. Para mostrar dicho resultado, realizaron un ejercicio del PIB de Estados Unidos con un modelo de tres estados. Sus resultados dejan un error de pronóstico de 164 periodos de los últimos datos menor que el modelo lineal. La diferencia entre los errores la comprueban con pruebas estadísticas frecuentistas de la literatura en pronóstico.

En el presente trabajo se aplica el enfoque de Boot y Pick (2017); es decir, las probabilidades suavizadas se sustituirán por pesos (w_t) óptimos para pronosticar $\hat{y}_{T+1} = X'_{T+1}\hat{\beta}(w)$. Donde los pesos, como se mencionó anteriormente, se eligen para minimizar el error de pronóstico en modelos de cambio de régimen (MSFE); esta metodología se denotará con OPT.

La Tabla 4 muestra la media del MSFE con base en el modelo lineal autorregresivo LAR, el pronóstico estándar de modelos de cambio de régimen MS y el óptimo OPT, para 64 periodos anteriores al último dato más 40 periodos de entrenamiento.¹² El resultado muestra que la media de MSFE del pronóstico OPT del IP (0.4422) es menor que el modelo lineal (0.4772). Incluso la media del MS también es menor (0.4428) que la del lineal. Por tanto, se puede concluir que la especificación sugerida por el factor de Bayes también tiene errores de pronóstico dentro de la muestra menores que el modelo lineal autorregresivo.

¹² Se utilizó el código de MATLAB disponible por los autores para estimar el MSFE.

Finalmente, en la parte inferior de la Tabla 4 presentamos el resultado para el pronóstico fuera de la muestra del IP. El IP en $T+1$ (es decir, abril de 2018) se estimó con el modelo lineal, el procedimiento estándar MS y aplicando los pesos óptimos (OPT). De esta estimación, se calculó el error al cuadrado con el dato real de 109.57 para el mes de abril. El menor error está dado cuando se aplican los pesos óptimos (0.0031), seguido por el de la metodología MS y tercero, el modelo lineal que presenta el error cuadrado más alto (0.0385) para este periodo de tiempo en específico.

Tabla 4
MSFE del IP y pronóstico dentro y fuera de la muestra

Modelo	LAR	MS	OPT
MSFE (media)	0.4772	0.4428	0.4422
Dato de Apr-2018			109.5700
Pronóstico (T+1)	109.7662	109.8329	109.6257
Error al cuadrado	0.0385	0.0692	0.0031

Nota: La estimación del MSFE está basada en Boot y Pick (2017).

Fuente: Elaboración propia.

En resumen, con la metodología sugerida por Boot y Pick (2017), determinamos que los errores de pronóstico del modelo MS el IP en México tiene menores errores que el modelo lineal autorregresivo lo que en la práctica es difícil de superar.

■ Conclusiones

En este trabajo, estudiamos la modelación econométrica y pronóstico dentro y fuera de la muestra, de procesos autorregresivos lineales y con parámetros markovianos sujetos a cambio de régimen de las tasas de crecimiento mensual del IP en México de 1980 a 2018. Con base en la metodología bayesiana, fue posible la estimación de la ML de cada modelo para llevar acabo la selección de estos a través del factor de Bayes de acuerdo con su ajuste a los datos. Esta metodología nos permitió elegir el número de rezagos, los parámetros sujetos a cambio de régimen y el número de estados posibles. Dado el modelo con mayor probabilidad marginal, se llevó a cabo un ejercicio de pronóstico para mostrar que la especificación markoviana (MS) tiene errores de pronóstico menores que el principal competidor, el modelo lineal autorregresivo.

Del análisis exploratorio de las tasas de crecimiento mensual del IP no hay ninguna evidencia de una mezcla de distribuciones normales, ya que la serie parece distribuirse con largas colas y media de cero. Sin embargo, se encontraron tres y cuatro estados con diferente media y volatilidad con base en el factor de Bayes, resultado similar a Frühwirth-Schnatter (2001). Sobre la especificación con más alta ML, se llevó a cabo un ejercicio de pronóstico en comparación del modelo lineal. Primeramente, se calcularon los pesos óptimos que minimizan el error de pronóstico en modelos MS basado en Boot y Pick (2017). Dados estos pesos, se prueba que la media de estos errores (dentro de la

muestra) es menor en el modelo MS que en lineal. En segundo lugar, aplicando los pesos óptimos y estimando el pronóstico fuera de la muestra para el mes de abril de 2018 contra la metodología estándar MS y para el modelo lineal, también se tiene un menor error cuadrado que los demás. Es decir, de este ejercicio, el modelo MS seleccionado por el factor de Bayes para el caso del IP tiene menor error de pronóstico dentro y fuera de la muestra que el modelo lineal.

Es importante mencionar que la metodología bayesiana sugerida en este trabajo resuelve problemas de no identificación de parámetros MS para llevar a cabo comparación entre modelos (*nuisance-parameters*), el número reducido de observaciones en algún estado no es problema mientras que la estimación por MLE falla. Lo anterior, porque las propiedades asintóticas son aplicadas al número de simulaciones más que en el tamaño de la muestra como en el enfoque de econometría frecuentista. Por último, bajo un adecuado muestreo de los parámetros todos los momentos pueden ser estimados con más precisión sin necesidad de la optimización como MLE, donde en caso de no identificación y pocas observaciones en algún estado, la matriz de información suele ser singular.

Finalmente, se deja para estudios posteriores el caso de series financieras de alta frecuencia con modelos de heterocedasticidad condicional (ARCH y GARCH) con parámetros de cambio de régimen markovianos (MS). Así mismo, este trabajo intenta aportar evidencia de que una posible especificación exógena de los estados podría estar mal especificada. Por ejemplo, en la mayoría de la literatura con datos de México citada anteriormente se imponen a priori dos estados. Por otro lado, es importante mencionar que para trabajos posteriores de co-movimientos, determinantes de la serie estudiada por modelos estructurales y VAR, así como las clásicas pruebas de estacionariedad y cointegración deberán tenerse en cuenta que los parámetros no son constantes en el tiempo, por lo que las implicaciones económicas de los resultados podrían ser distintas.

Otra línea de investigación futura podría extenderse considerando la interacción de la producción industrial de México con la de Estados Unidos, tomando en cuenta su sincronía, efectos de transmisión y los resultados de este trabajo.¹³

■ Bibliografía

- Andrews, D. & W. Ploberger (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62 (6), 1383-1414.
- Bauwens, L., Dufays, A. & Rombouts, J. V. (2014). Marginal likelihood for Markov-switching and change-point (GARCH) models. *Journal of Econometrics*, 178 (parte 3), 508-522.
- Bazdresch, S. & Werner, A. (2005). Regime switching models for the Mexican peso. *Journal of International Economics*, 65 (1), 185-201.
- Benavides, G. & Capistrán, C. (2012). Forecasting exchange rate volatility: The superior performance of conditional combinations of time series and option implied forecasts. *Journal of Empirical Finance*, 19 (5), 627-639.

¹³ Véase De León (2004).

- Boot, T. & Pick, A. (2017). Optimal forecasts from Markov switching models. *Journal of Business & Economic Statistics*, 36 (4), 1-15.
- Cabrera, G., Coronado, S., Rojas, O. & Romero-Meza, R. (2018). A Bayesian approach to model changes in volatility in the Mexican stock exchange index. *Applied Economics*, 50 (15), 1716-1724.
- Carrasco, M., Hu, L. & Ploberger, W. (2014). Optimal test for Markov switching parameters. *Econometrica*, 82 (2), 765-784.
- Clements, M. P. & Krolzig, H. M. (1998). A comparison of the forecast performance of Markov-switching y threshold autoregressive models of us GNP. *The Econometrics Journal*, 1 (1), 47-75.
- De León Arias, A. (2004). El crecimiento de México y su interdependencia con Estados Unidos. *Comercio Exterior*, 54 (7), 612-619.
- Diebold, F. & Rudebusch, G. (1999). *Business cycles: Durations, dynamics, and forecasting*. New Jersey, Estados Unidos: Princeton University Press.
- Engel, C. (1994). Can the Markov switching model forecast exchange rates? *Journal of International Economics*, 36 (1), 151-165.
- Engel, C. & Kim, C. J. (1999). The long-run U.S./U.K. real exchange rate. *Journal of Money, Credit and Banking*, 31 (3), 335-356.
- Frühwirth-Schnatter, S. (2001). Fully Bayesian analysis of switching gaussian state space models. *Annals of the Institute of Statistical Mathematics*, 53 (1), 31-49.
- Frühwirth-Schnatter, S. (2004). Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques. *Econometrics Journal*, 7 (1), 143-167.
- Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models* (Springer Series in Statistics). Nueva York, Estados Unidos: Springer.
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series y the business cycle. *Econometrica*, 57 (2), 357-384.
- Hamilton, J. D. (1994). *The time series analysis*. Estados Unidos: Princeton University Press.
- Hamilton, J. D. (2005). What's real about the business cycle? Working Paper 11161. Estados Unidos: National Bureau of Economic Research.
- Hamilton, J. D. & Raj, B. (2002). *Advances in Markov-switching models: Applications in business cycle research and finance*. Estados Unidos: Springer.
- Hamilton, J. D. & Susmel, R. (1994). Autoregressive conditional heteroskedasticity y changes in regime. *Journal of Econometrics*, 64 (1), 307-333.
- Hansen, B. (1996). Erratum: The likelihood ratio test under nonstard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics*, 11 (2), 195-198.
- Heath, E. B. & Kopchak, S. J. (2015). The response of the Mexican equity market to us monetary surprises. *Journal of Emerging Market Finance*, 14 (2), 87-111.
- Islas-Camargo, A., Cortez, W. W. & Sanabria Flores, T. P. (2017). Is Mexico's forward exchange rate market efficient? *Revista Mexicana de Economía y Finanzas*, 13 (2), 247-272.

- Kass, R. E. & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90 (430), 773-795.
- Kaufmann, S. & Frühwirth-Schnatter, S. (2002). Bayesian analysis of switching arch models. *Journal of Time Series Analysis*, 23 (4), 425-458.
- Kim, C. J. & Nelson, C. R. (1999a). Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *The Review of Economics and Statistics*, 81 (4), 608-616.
- Kim, C. J. & Nelson, C. R. (1999b). State-Space models with regime switching: Classical and Gibbs-sampling approaches with applications (Volumen 1). Boston, Estados Unidos: MIT Press.
- López-Herrera, F. (2004). Modelado de la volatilidad y pronóstico del índice de precios y cotizaciones de la bolsa mexicana de valores. *Contaduría y Administración* (213), 43-72.
- López-Herrera, F. & Venegas-Martínez, F. (2011). *Métodos no lineales en series económicas y financieras. Modelado de la volatilidad del mercado mundial de capitales durante la crisis mundial mediante una cadena de Markov*. Guadalajara, México: Universidad de Guadalajara.
- López-Herrera, F., Ortiz-Arango, F. & Venegas-Martínez, F. (2011). *Crecimiento y desarrollo económico en México*. Modelado de la volatilidad del índice de precios y cotizaciones de la Bolsa Mexicana de Valores con cambios markovianos de régimen. México: Benemérita Universidad Autónoma de Puebla.
- Mejía-Reyes, P. & Díaz-Carreño, M. A. (2010). Búsqueda de asimetrías a lo largo del ciclo económico de México. *Investigación Económica*, 69 (271), 15-42.
- Mejía-Reyes, P., Martínez Gómez, J. A. & Rendón Balboa, W. L. (2005). Ciclos económicos clásicos en la producción industrial de México. *Investigación Económica*, 64 (254), 91-124.
- Meng, X. I. & Wong, W. H. (1996). Simulating ratios of normalizing constants via a simple identity: A theoretical exploration. *Statistica Sinica*, 6 (4), 831-860.
- Timmermann, A. (2000). Moments of Markov switching models. *Journal of Econometrics*, 96 (1), 75-111.