Cámara Chilena de la Construcción A.G.
Gerencia de Estudios

Nº 73

Marzo 2013

Evaluación de modelos de predicción para la venta de viviendas

Fco. Javier Lozano

Evaluación de modelos de predicción para la venta de viviendas*

Fco. Javier Lozano**

Marzo 2013

Resumen

Este documento de trabajo evalúa la capacidad predictiva de diferentes modelos econométricos en horizontes de predicción de tres, seis y doce meses, con el objetivo de mejorar los pronósticos sobre venta de viviendas que publica la Cámara Chilena de la Construcción. Para ello, se estiman cinco tipos distintos de modelos, entre los que destacan los vectores autorregresivos de tipo bayesiano, que han recibido amplia aceptación en la última década. El principal resultado obtenido es que, en la mayoría de casos, los modelos BVAR entregan predicciones más precisas que los modelos clásicos, lo cual es coherente con la evidencia encontrada en diversas aplicaciones macroeconómicas y sectoriales de este tipo de modelos.

* El autor agradece los valiosos comentarios del equipo de la Gerencia de Estudios de la Cámara Chilena de la Construcción A.G. Cualquier error es responsabilidad del autor.
** Gerencia de Estudios, Cámara Chilena de la Construcción A.G. E-Mail: flozano@cchc.cl

1. Introducción

El sector inmobiliario afecta a la macroeconomía a través de múltiples canales, involucrando a gran cantidad de agentes económicos. Por un lado, están las empresas constructoras e inmobiliarias, que generan empleo y actividad económica a través de la construcción y venta de viviendas. Por otro, están los hogares, que destinan una porción significativa de su renta a la compra o arriendo de vivienda. Entre medio, las instituciones financieras, encargadas de proveer financiación tanto a empresas como a hogares.

En los últimos años los Bancos Centrales han prestado mayor atención a la evolución del mercado inmobiliario, sobre todo teniendo en cuenta el rol que este sector tuvo en la gestación de la crisis financiera de 2007 en Estados Unidos y su posterior contagio a otras zonas. Así, el Banco Central de Chile, en los dos últimos *Informes de Estabilidad Financiera*, ha dedicado mayores esfuerzos por entender las dinámicas del sector y sus efectos sobre el ciclo económico. Igualmente, la Reserva Federal de Estados Unidos hace referencia, en casi todos sus comunicados oficiales de coyuntura económica, al sector inmobiliario por la importancia que este tiene sobre el resto de sectores en términos de empleo y actividad económica. Incluso hay autores que han considerado que el sector inmobiliario "es el ciclo económico" (Leamer, 2007), dando a entender que la inversión residencial es el mejor indicador para prever un cambio en el ciclo.

El objetivo de este documento de trabajo es evaluar diversos modelos econométricos para la predicción de la venta de viviendas en horizontes de predicción de tres, seis y doce meses, de tal manera que sea posible contar con pronósticos lo más precisos posible sobre la evolución del sector inmobiliario. Para la Cámara Chilena de la Construcción este asunto no es trivial, pues una parte importante de la actividad del gremio está en directa relación con la construcción y venta de viviendas. El aporte fundamental de este trabajo radica en la necesidad de entregar predicciones más certeras de la actividad inmobiliaria residencial, las cuales se publican con periodicidad cuatrimestral en el Informe MACh (Macroeconomía y Construcción).

Hasta el momento, la predicción de venta de viviendas se realizaba empleando un modelo de tipo autorregresivo con indicadores adelantados, en base a la metodología propuesta por Demers (2005). Si bien este tipo de modelos ha mostrado un buen desempeño en su capacidad predictiva al compararlo con modelos estructurales y univariantes, existe la convicción de que los pronósticos pueden ser mejorados al adoptar una metodología basada en restricciones de tipo bayesiano, al estilo de Litterman (1979) y teniendo en cuenta otros aportes posteriores. La aplicación de técnicas bayesianas para la mejora de las predicciones acumula más de 30 años de desarrollo metodológico, siendo este especialmente significativo en la última década gracias a los avances informáticos que han permitido estimar modelos a gran escala.

Este documento está organizado en cinco secciones. La primera presenta la introducción al tema tratado. La segunda expone los distintos modelos considerados para la evaluación de las predicciones. La tercera describe la manera en que cada modelo es estimado. La cuarta sección analiza los principales resultados obtenidos, comparando la precisión de cada modelo empleado. Finalmente, la quinta sección presenta las conclusiones.

2. Modelos de predicción

2.1. Modelos ARIMA

Los modelos más sencillos para la predicción de series temporales son los univariantes, por depender estos únicamente de valores pasados de la propia serie para explicar la dinámica de la variable de interés. Esto supone, al mismo tiempo, una ventaja y un inconveniente, pues si bien se trata de modelos fáciles de especificar y estimar, pierden capacidad de análisis al no incorporar otras variables explicativas.

Dentro de los modelos univariantes para series temporales, los denominados ARIMA (modelos autorregresivos integrados de medias móviles) han sido los más extensamente empleados para predicción. Box y Jenkins (1970) desarrollaron la metodología para la identificación, estimación y diagnóstico de estos modelos dinámicos.

El modelo autorregresivo se define como aquel en que la variable endógena en un período t es explicada por las observaciones de ella misma correspondientes a periodos anteriores, incluyendo un término de error que se asume ruido blanco. La especificación general de un modelo autorregresivo de orden p (*AR[p]*) es la siguiente:

$$Y_t = \rho_0 + \rho_1 Y_{t-1} + \rho_2 Y_{t-2} + \cdots + \rho_p Y_{t-p} + \varepsilon_t \qquad (1)$$

El modelo denominado de medias móviles se define como aquel que explica el valor de una variable en un periodo t en función de una sucesión de errores correspondientes a periodos anteriores, con su respectiva ponderación. La especificación de un modelo de medias móviles de orden q (*MA[q]*) es como sigue:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \qquad (2)$$

Por tanto, la combinación de modelos autorregresivos y de medias móviles resulta en la siguiente especificación del modelo *ARMA(p,q)*:

$$Y_t = \rho_0 + \rho_1 Y_{t-1} + \cdots + \rho_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \qquad (3)$$

La aplicación de este tipo de modelos en ejercicios de estimación y predicción de series temporales ha sido generalizada durante décadas. Además, los modelos ARIMA son generalmente utilizados como referencia para la evaluación de predicciones realizadas con modelos más avanzados. Johnston y DiNardo (1997) mostraron cómo emplear esta metodología para estimar y generar pronósticos para la serie mensual de inicios de obras de viviendas en Estados Unidos. Para ello especificaron modelos *AR(1)* y *ARMA(1,1)* y compararon las predicciones entregadas por cada uno. Así, determinaron que la inclusión del componente de medias móviles suponía una mejora en las predicciones en comparación con el modelo puramente autorregresivo.

2.2. Modelos ARLI

Como una manera de superar la limitación de los modelos univariantes descritos anteriormente, en los años ochenta un grupo de académicos consideró necesario complementar los modelos ARIMA con la inclusión de determinadas variables que pudieran servir como indicadores adelantados. De esta manera surgieron los modelos denominados ARLI, es decir, autorregresivos con indicadores adelantados.

La especificación general del modelo ARLI es como sigue:

$$Y_t = \sum_{i=1}^{p} \rho_i Y_{t-i} + \sum_{j=1}^{q} \varphi_j X_{t-j} + \varepsilon_t \tag{4}$$

Donde Y es la variable dependiente, de la cual se quieren obtener predicciones; X es un conjunto de variables explicativas que sirven como indicadores adelantados para la variable explicada; ε es el término de error; y el orden de rezagos es p y q. Por tanto, el primer elemento de la parte derecha de la ecuación (4) es el componente autorregresivo, mientras que el segundo elemento es el componente de indicadores adelantados.

Garcia-Ferrer et al. (1987) realizaron el ejercicio de predecir la tasa de crecimiento económico de un grupo de nueve países usando un modelo autorregresivo de orden 3. Al compararlo con las predicciones obtenidas de modelos *naive*, vieron que en ambos casos las proyecciones eran malas. Por esta razón decidieron incorporar como variables explicativas la rentabilidad bursátil (primer y segundo rezago) y el crecimiento de la masa monetaria (primer rezago) de cada país. Con esto consiguieron reducir el error de predicción, mejorando los pronósticos obtenidos con modelos *naive* y autorregresivos.

Posteriormente Zellner y Hong (1989) y Zellner y Tobias (2000) aplicaron esta misma metodología para la predicción de la tasa de crecimiento económico en un grupo formado por dieciocho países. Sus resultados confirmaron que los modelos ARLI entregan proyecciones más exactas –con menor error cuadrático medio– que modelos *naive* y autorregresivos. Sin embargo, al desarrollar ciertas técnicas más avanzadas de predicción (desagregación y contracción de tipo bayesiano) consiguieron mejorar las predicciones de los modelos autorregresivos con indicadores adelantados.

Con respecto al sector inmobiliario, la única aplicación de modelos ARLI encontrada corresponde a Demers (2005), quien comparó la precisión de las predicciones obtenidas de modelos de tipo estructural con aquellas entregadas por los modelos ARLI. Estos últimos proporcionaron pronósticos más exactos de la inversión residencial, reduciendo el error cuadrático medio hasta casi la mitad en comparación con los modelos estructurales. En general, los modelos de tipo estructural presentan malos resultados en la predicción porque requieren estimar un número elevado de parámetros.

2.3. Modelos VAR

Una de las herramientas econométricas con mayor desarrollo en los últimos veinte años es el vector autorregresivo (VAR). Su uso se ha generalizado en las aplicaciones macroeconómicas, en gran parte debido a lo sencillo que resulta su formulación y estimación, pero también por los buenos resultados que ofrece en términos de análisis de impacto y predicción.

La introducción de los vectores autorregresivos se debe a Sims (1980), quien hizo una dura crítica a los grandes modelos macroeconómicos, basados en ecuaciones simultáneas, que hasta ese momento se empleaban para predicción y análisis de impacto. La crítica de Sims se fundamentó en el hecho de que estos modelos estaban basados en restricciones poco creíbles y sin soporte en la teoría económica. De esta manera, Sims consideró factible estimar grandes modelos macroeconómicos sin imponer restricciones *a priori* y tratando a todas las variables como endógenas. La única restricción necesaria para estimar este tipo de modelos sería determinar el orden de rezagos.

La especificación genérica de un modelo VAR es la siguiente:

$$Y_t = C + \sum_{i=1}^{p} A_i\, Y_{t-i} + \varepsilon_t \tag{5}$$

Donde Y es un vector de variables; A es la matriz de coeficientes asociados a las variables; ε es el término de error; y p es el orden de rezagos. Por lo general se asume que el término de error es un proceso de ruido blanco, con media cero y matriz de covarianzas positiva y constante en el tiempo.

Posterior a la introducción de los vectores autorregresivos, comenzó a desarrollarse la teoría de la cointegración a partir del aporte inicial de Engle y Granger (1987). En general, las variables utilizadas para la estimación de modelos VAR deberían ser estacionarias, lo cual implica que su media y su varianza son constantes en el tiempo. Esto no siempre ocurre con las series temporales, pues pueden presentar tendencias o componentes estacionales que son incompatibles con la estacionariedad; en estos casos, es frecuente recurrir a la diferenciación de la serie para transformarla en estacionaria. Engle y Granger (1987) demostraron que la combinación lineal de series integradas del mismo orden (d) podría, en determinados casos, tener un orden de integración inferior (d-b, con b>0). El ejemplo más recurrente es cuando se tienen dos series X_t e Y_t, ambas integradas de orden 1 (no estacionarias) y su combinación lineal $Z_t = X_t - aY_t$ es una serie integrada de orden 0. Esto significa que la diferencia entre la series X e Y es estacionaria, lo que implica que su evolución en el tiempo será estable.

Cuando las variables del modelo VAR son cointegradas, el procedimiento tradicional de tomar diferencias para convertirlas en estacionarias es erróneo, pues en ese caso se estaría ignorando la relación que las series mantienen en el largo plazo. En estos casos, lo correcto sería estimar un modelo de vectores de corrección del error (VECM), el cual incluye las relaciones de cointegración que existen entre las variables empleadas.

La especificación del modelo VECM es como sigue:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \theta_i \, \Delta Y_{t-i} + \varepsilon_t \qquad (6)$$

Donde ΔY representa la primera diferencia del conjunto de variables de interés; ΠY_{t-1} es el componente de largo plazo; y θ_i son los parámetros de corto plazo.

Como ya se mencionó anteriormente, el empleo de modelos VAR ha sido generalizado en las últimas décadas, por su facilidad de uso y los buenos resultados que ofrece. Las aplicaciones más frecuentes se han generado en el ámbito de la macroeconomía y han estado centradas tanto en análisis de impacto como en predicción. También ha existido una aplicación generalizada en economía sectorial y, en algunos casos, en temas microeconómicos. Con respecto al sector inmobiliario, Diebold (2007) mostró cómo una especificación correcta de un modelo VAR para los inicios y términos de construcción de viviendas podía entregar predicciones bastante precisas, incluso anticipando puntos de inflexión en las variables de interés.

2.4. Modelos BVAR

Si bien los modelos VAR tuvieron, y siguen teniendo, una amplia aceptación en las aplicaciones econométricas, desde el principio se reconoció que su principal debilidad provenía del elevado número de parámetros a estimar. El problema es doble: en primer lugar, la inclusión de más variables en el modelo supone estimar más parámetros, lo cual conlleva una pérdida significativa en los grados de libertad y afecta negativamente a la estimación del modelo y las predicciones obtenidas; esto ocurre también al incluir más rezagos en el modelo (Canova, 2007; Litterman, 1979). En segundo lugar, para solventar la falta de grados de libertad, se suele recurrir a la no inclusión de determinadas variables o número de rezagos, lo cual implica imponer restricciones de exclusión y puede generar sesgos por omisión de variables (Banbura *et al*., 2010; Todd, 1988a).

Para superar esta debilidad de los modelos VAR, una de las alternativas propuestas que ha recibido mayor atención es la aplicación de técnicas bayesianas para la estimación de vectores autorregresivos, lo cual da origen a los modelos BVAR. Las primeras propuestas en este sentido surgen en la Reserva Federal de Minneapolis y en la Universidad de Minnesota: Sims (1980) reconoció que los vectores autorregresivos podrían mejorar sus pronósticos incorporando métodos bayesianos y Litterman (1979) elaboró la primera familia de *priors* (distribución de probabilidad *a priori*), comúnmente conocida como "*prior* de Minnesota".

Posteriormente se fueron generando distribuciones de probabilidad *a priori* más avanzadas, con la intención de superar las limitaciones impuestas por la *prior* de Minnesota. En los últimos años, la utilización de modelos BVAR, tanto para predicción como para análisis de impacto, ha recibido especial atención gracias a los avances informáticos que han permitido la estimación de grandes modelos que incorporan técnicas bayesianas para su estimación. De esta manera, varios Banco Centrales, entre los que destacan la Reserva Federal de Estados Unidos y el Banco

Central Europeo, han evaluado la precisión de los modelos BVAR y los han incorporado como herramienta clave para generar pronósticos de variables macroeconómicas.

La aplicación de técnicas bayesianas en la estimación de vectores autorregresivos presenta diversas ventajas. En general, la especificación de modelos BVAR permite superar la escasez de grados de libertad típica de los vectores autorregresivos, sin tener que recurrir a restricciones de exclusión. Se trata de un enfoque distinto a la estimación clásica: incluir tantas variables y rezagos como se considere necesario, especificando un conjunto de creencias *a priori* sobre los parámetros para complementar los datos (Litterman, 1979; Todd, 1988a). Estas creencias *a priori* se especifican como una distribución de probabilidad sobre los parámetros del modelo, de tal manera que no sería necesario excluir determinadas variables o rezagos, sino que se considerarían como variables aleatorias centradas en cero y con varianza decreciente (Ciccarelli y Rebucci, 2003). En caso de que este supuesto sobre la distribución de probabilidad de los parámetros no se cumpla y existan relaciones significativas entre variables y rezagos que *a priori* se consideraron menos relevantes, los propios datos anularían este supuesto (Dua *et al.,* 2006). De esta manera, al aplicar técnicas bayesianas en la estimación de vectores autorregresivos es posible incorporar mayor interacción en el modelo, al mismo tiempo que se consigue mayor precisión si la distribución de probabilidad *a priori* de los parámetros es la correcta (Keller, 2007; Litterman, 1979).

La econometría bayesiana se basa en el Teorema de Bayes, según el cual la probabilidad condicional de un evento aleatorio A dado B puede ser expresada en términos de la distribución de probabilidad condicional del evento B dado A y la distribución de probabilidad marginal de b:

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)} \tag{7}$$

Siguiendo a Koops (2003), la aplicación de esta regla de probabilidad al modelo de regresión general sería como sigue: dado un vector o matriz de datos (y) y un vector o matriz que contiene los parámetros del modelo (θ), el interés se centra en aprender de θ en base a los datos y. Esto se puede expresar como:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \tag{8}$$

Dado que el interés se centra en los parámetros (θ) y no en los datos, es posible ignorar el término $p(y)$ y reescribir la ecuación (8) como:

$$p(\theta|y) \propto p(y|\theta)p(\theta) \tag{9}$$

El término $p(\theta|y)$ se denomina *densidad posterior* y contiene todo lo que se sabe sobre los parámetros una vez que se han visto los datos; $p(y|\theta)$ es la función de verosimilitud y representa la función de densidad de probabilidad de los datos dados los parámetros; y $p(\theta)$ es la *densidad priori*, independiente de los datos, y que contiene todo lo que se sabe sobre los parámetros antes de haber visto los datos. Por tanto, de la expresión (9) se entiende que la *posterior* es proporcional a la verosimilitud por la *priori*.

Como se mencionó anteriormente, Litterman (1979) propuso la primera familia de *prioris*, la denominada "*priori* de Minnesota". Esta especificación sugiere que el paseo aleatorio con deriva es una aproximación razonable al comportamiento de muchas series económicas. Por tanto, la distribución *a priori* de los parámetros del vector autorregresivo estaría centrada alrededor de la siguiente ecuación:

$$Y_t = c + Y_{t-1} + \varepsilon_t \tag{10}$$

En términos algebraicos, a partir de la especificación genérica de un vector autorregresivo de la forma:

$$Y_t = A_0 + A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \varepsilon_t, \varepsilon_t \sim N(0, \Sigma) \tag{11}$$

Los parámetros de interés serían A_k y Σ. Una de las características que hacen tan simple a la formulación de Litterman es que la varianza del error (Σ) se asume diagonal y se reemplaza por una estimación. De esta manera, la estimación del modelo BVAR se reduce a aplicar una distribución *prior* para los parámetros contenidos en A_k. En este sentido, la *prior* de Minnesota considera que estos parámetros se distribuyen según una Normal, con media cero (salvo en el primer rezago de cada variable dependiente) y matriz de covarianzas diagonal. Los elementos de esta diagonal se definen como:

$$v_{ij,l} = \begin{cases} a_1/p^2 & \text{para los coeficientes de rezagos de la propia variable} \\ (a_2\sigma_i)/(p^2\sigma_j) & \text{para los coeficientes de rezagos de otras variables} \\ a_3\sigma_i & \text{para los coeficientes de variables exógenas} \end{cases} \tag{12}$$

Donde σ_i es el i-ésimo elemento diagonal de la matriz de covarianzas Σ.

De esta manera, se asume que los parámetros del modelo tienen *a priori* media cero, excepto los coeficientes del primer rezago que tienen media uno. Esto supone contraer los elementos de la diagonal de la matriz A en la ecuación (11) hacia uno y los otros coeficientes fuera de la diagonal hacia cero. De esta manera, se reduce significativamente el riesgo de sobre-parametrizar el modelo (Koop y Korobilis, 2009). Además, se asume que la varianza de los parámetros es decreciente con el número de rezagos y con los rezagos de otras variables. Siguiendo a Ciccarelli y Rebucci (2003), esto supone admitir que las series económicas cumplen tres regularidades estadísticas: comportamiento tendencial, mayor impacto de los rezagos más recientes y mayor impacto de los rezagos propios frente a los de otras variables.

Por tanto, la estimación del modelo BVAR haciendo uso de la *prior* de Minnesota se reduce a la selección de tres hiper-parámetros: a_1, a_2 y a_3, que determinan el comportamiento de la varianza de los parámetros del modelo. La especificación propuesta, como se mencionó previamente, introduce interesantes propiedades en la dinámica del modelo. Por un lado, conforme aumenta el número de rezagos, los coeficientes se contraen hacia cero, es decir, pierden importancia. Por otro, se da mayor peso a los rezagos de la propia variable frente al resto de variables.

Diversos autores han propuesto mejoras para la *prior* de Minnesota. Entre estos, los aportes más destacados han sido los de Doan, Litterman y Sims (1983) y Sims y Zha (1996). En el primer caso, la *prior* de Minnesota se complementa con una *prior* sobre la suma de los coeficientes del modelo, la cual se implementa a través de un hiper-parámetro que controla el grado de ajuste de la suma de coeficientes hacia cero. Teniendo en cuenta que un vector autorregresivo con corrección de errores (6) se puede escribir como:

$$\Delta Y_t = c - \left(I_n - A_1 - \cdots - A_p\right) Y_{t-1} + B_1 \Delta Y_{t-1} + \cdots + B_{p-1} \Delta Y_{t-p+1} + \varepsilon_t \quad (13)$$

La *prior* propuesta por Doan *et al.* (1983) implica que la suma de coeficientes $(I_n - A_1 - \ldots - A_p)$ está centrada en uno para los rezagos de la propia variable, y en cero para los rezagos de las otras variables. Además introduce correlación entre los coeficientes de cada variable en cada ecuación del modelo (Giannone, 2012). Esta restricción implicaría la presencia de una raíz unitaria en cada ecuación y descartaría la existencia de relaciones de cointegración.

Con respecto al aporte de Sims y Zha (1996), este consiste en la inclusión de hiper-parámetros adicionales a la *prior* de Minnesota, los cuales permiten considerar dentro del modelo comportamientos tendenciales y relaciones de cointegración entre las variables. Partiendo de la especificación propuesta por Litterman, Sims y Zha definen la varianza del coeficiente correspondiente al rezago *p* de la variable *i* en la ecuación *j* como:

$$V_{0p,ij} = \left(\frac{\lambda_0 \lambda_1}{\sigma_j p^{\lambda_3}}\right)^2 \quad (14)$$

Donde λ_0 controla el ajuste general de la *prior*, λ_1 gobierna el ajuste en torno al componente autorregresivo y λ_3 controla la contracción de la varianza conforme aumenta el número de rezagos. Una de las diferencias más significativas con respecto a la especificación de Litterman se encuentra en que la *prior* Sims-Zha no incorpora la distinción entre rezagos propios y rezagos de otras variables. Esto se debe a que el modelo empleado para especificar la *prior* es de ecuaciones simultáneas, por tanto no tiene sentido hablar de variables dependientes.

La *prior* Sims-Zha agrega, además, dos hiper-parámetros que permiten incorporar en el modelo comportamientos dinámicos más complejos. El primero de ellos es el propuesto por Doan et *al.* (1983) sobre la suma de los coeficientes del modelo, el cual se implementa a través del parámetro μ_5 que controla el grado de ajuste de la suma de coeficientes hacia cero. Este hiper-parámetro permite incorporar en el modelo la presencia de raíces unitarias y tendencias lineales. El segundo hiper-parámetro propuesto (μ_6) facilita la incorporación de relaciones de cointegración entre las variables del modelo.

Uno de los aspectos más criticados de la *prior* de Minnesota y sus modificaciones posteriores es el hecho de que la matriz de covarianzas del error ($\sum$) se considera diagonal y se reemplaza por una estimación. Esto conlleva no tener en cuenta la posibilidad de correlación entre los residuos de diferentes ecuaciones (Banbura et *al.*, 2010), además de ignorar cualquier opción de incertidumbre sobre dicho componente (Koop y Korobilis, 2009).

Para solventar este obstáculo, una de las alternativas que ha recibido más aceptación es la de Kadiyala y Karlsson (1997), quienes relajaron el supuesto de matriz de covarianzas diagonal y fija utilizando una distribución *prior* de tipo conjugada. De este modo, la denominada *prior* Normal-Wishart considera que los parámetros del modelo se distribuyen de acuerdo a las siguientes especificaciones:

$$A|\sum \sim N(B_0, V_0) \quad (15)$$

$$\sum \sim IW(v_0, S_0) \quad (16)$$

Donde la matriz de coeficientes (B_0) se considera diagonal, de tal manera que los coeficientes tienen *a priori* media cero, salvo los del primer rezago de cada variable que tienen media uno.

Mientras que la matriz de covarianzas (V_0) está compuesta por elementos definidos como en (14). Por su parte, la matriz de covarianzas del error tiene media $v_0=m+1$ (con m igual al número de variables en el modelo) y varianza S_0 función del hiper-parámetro λ_0.

Por tanto, de acuerdo a esta *prior*, la distribución *a priori* de los coeficientes es de tipo Normal, mientras que la distribución de la matriz de covarianzas del error es de tipo Wishart inversa. Esto permite considerar en el modelo la existencia de correlación entre las distintas ecuaciones (Robertson y Tallman, 1999).

Como ya se mencionó anteriormente, la aplicación de métodos bayesianos para la estimación de vectores autorregresivos ha recibido gran interés desde sus comienzos. Este interés tuvo cierto apogeo en los últimos años, gracias a las modificaciones propuestas para su mejora, así como por los avances informáticos que han facilitado su aplicación. En general, los modelos BVAR han sido empleados con la finalidad de mejorar los pronósticos acerca de variables macroeconómicas de interés. Así, Doan *et al.* (1983), Litterman (1986) y Todd (1988a), entre otros, reportan mejoras sustanciales en la precisión de las proyecciones al imponer distribuciones priori sobre los coeficientes del vector autorregresivo, de tal manera que el error cuadrático medio es menor comparado con modelos univariantes y vectores autorregresivos sin restricciones. Además, suele ocurrir que la precisión de los modelos BVAR no se ve afectada al incrementar el tamaño del modelo, ya sea al incluir más variables o más rezagos, lo cual es una significativa ventaja frente al problema de sobre-parametrización de los vectores autorregresivos (Giannone *et al.*, 2012).

Existen en la literatura internacional múltiples aplicaciones de modelos BVAR para la estimación y predicción de variables de tipo macroeconómico. Algunos ejemplos son: Banbura et al. (2010), quienes estimaron un modelo BVAR a gran escala (hasta 131 variables) para predecir el desempleo, la inflación y la tasa de interés de Estados Unidos; Gupta et al. (2009) plantearon diversas especificaciones para la prior de su vector autorregresivo, con el objetivo de mejorar los pronósticos de variables macroeconómicas de Sudáfrica, obteniendo como resultado que una prior más ajustada entrega mejores predicciones que una prior más flexible; Llosa et al. (2006) quienes estimaron un modelo BVAR a partir de la especificación de Litterman para predecir la inflación en Perú; Kenny et al. (1998) también buscaron mejorar los pronósticos para la inflación, en este caso de Irlanda, haciendo uso de la prior de Minnesota; Crone y McLaughlin (1999) aplicaron un modelo BVAR para la predicción de varias variables de interés económico en el área metropolitana de Philadelphia; Servinç y Ergün (2009) especificaron cinco modelos BVAR distintos, cada uno haciendo uso de una prior distinta, con el objetivo de mejorar las predicciones para el desempleo y la producción industrial de Turquía. En todos estos casos se repite un mismo fenómeno: los pronósticos entregados por los modelos BVAR, bajo diferentes especificaciones, mejoran significativamente las predicciones obtenidas a partir de modelos univariantes y vectores autorregresivos sin restricciones.

Aunque no hayan sido tan recurrentes en la literatura sobre modelos BVAR, se pueden encontrar algunas aplicaciones de esta metodología para el cálculo de funciones de respuesta al impulso (FRI). Por ejemplo, Ciccarelli y Rebucci (2003) estimaron un modelo de economía abierta para analizar el impacto de la política monetaria en la Eurozona, considerando las principales economías de dicha área: Alemania, Francia, Italia y España. Banbura et al. (2010), a partir de su modelo BVAR a gran escala, analizaron el efecto de un *shock* monetario sobre diversas variables económicas de interés en Estados Unidos. Joiner (2001) estimó un modelo de economía abierta para Australia y analizó la manera en que un *shock* en determinadas variables monetarias se distribuye en la economía, haciendo uso para ello de la *prior* Sims-Zha.

La aplicación de la metodología de vectores autorregresivos de tipo bayesiano para analizar la economía de Chile es escasa, pues solo se encontraron tres aplicaciones. La primera corresponde a Todd y Morandé (1988b), quienes aplicaron la *prior* de Minnesota con ciertas modificaciones para obtener predicciones de diversas variables de interés; sus resultados avalan que los modelos BVAR son capaces de reducir el error cuadrático medio de la predicción, aunque en términos absolutos este error es elevado. Las otras dos aplicaciones surgieron en los últimos años. Una corresponde a Jaramillo (2009), quien estudió los mecanismos de transmisión de la política monetaria y obtuvo proyecciones para las principales variables macroeconómicas a partir de modelos BVAR basados en la *prior* Sims-Zha, obteniendo como resultado que el desempeño de estos modelos era mejor que el de modelos tradicionales. La otra aplicación de vectores autorregresivos bayesianos al análisis de la economía chilena es de González (2012), que desarrolló un BVAR a gran escala, con más de 100 variables, para evaluar la capacidad predictiva y estudiar el efecto de *shocks* de tipo monetario y sectorial.

Con respecto al sector inmobiliario, solo se encontraron dos aplicaciones de modelos BVAR. Dua *et al.* (1996) estimaron diversos vectores autorregresivos para la venta de viviendas en Estados Unidos, haciendo uso de la *prior* de Minnesota. En general, encuentran que los modelos BVAR mejoran las predicciones entregadas por los vectores autorregresivos tradicionales, lo cual es más evidente en aquellos BVAR con mayor número de rezagos. La otra aplicación referida al sector inmobiliario corresponde a Gupta *et al.* (2009), quienes estimaron un VAR bayesiano con 143 variables para analizar el impacto de la política monetaria sobre diversas variables del sector para Estados Unidos.

3. Estimación

Para la estimación de los cuatro tipos de modelos se emplean las mismas variables. Como variable dependiente, de la cual se desea obtener predicciones, se considera la venta de viviendas en el Gran Santiago, medida en unidades. Como explicativas se consideran la tasa de crecimiento anual del IMACEC, la inflación anual y la tasa de política monetaria (TPM). La frecuencia de estas series es mensual, desde enero de 1994 hasta junio de 2011. Para la evaluación de las predicciones se usa el periodo que va de julio de 2011 a junio de 2012, es decir 12 meses.

Si bien podría haberse planteado una ecuación de demanda por vivienda más completa, se estima que la inclusión de la inflación, el IMACEC y la tasa de política monetaria (TPM) recoge en gran medida los principales fundamentos de la evolución de la economía. De hecho, los coeficientes de correlación entre las variables son relativamente elevados: 0,4 para TPM, 0,26 para IMACEC y 0,06 para inflación, lo cual evidencia cierta relación entre las variables. Además, la elección de estas tres variables se justifica por el hecho de que la predicción del modelo ARLI requiere contar con valores futuros de las variables explicativas, y en este caso, a través de las encuestas de expectativas del Banco Central,es posible contar con dichos datos.

La información sobre venta de viviendas se obtiene de las bases de datos de la Cámara Chilena de la Construcción. Para las variables explicativas (IMACEC, Inflación y TPM) se recurre a las bases del Banco Central de Chile. En los tres casos es necesario realizar algún empalme para poder conseguir una serie completa desde 1994. La tasa de política monetaria comenzó a utilizarse en mayo de 1995, por tanto los datos anteriores a esa fecha se obtienen a partir de la tasa real de los PRBC a 90 días. Para la inflación se utiliza la serie empalmada disponible en el Banco Central. Mientras que el IMACEC se completa usando las bases de referencia 2008, 1996 y 1986 que publica el Banco Central.

Antes de realizar las estimaciones de los modelos, se tiene en cuenta que, generalmente, las series económicas presentan comportamientos tendenciales o estacionales. Por esa razón, la serie de venta de viviendas se desestacionaliza mediante el método Census X12. Con respecto a la presencia de raíces unitarias en las series, se llevan a cabo las pruebas de Dickey-Fuller aumentada y GLS (Tablas 3 y 4 del Anexo). Los resultados son mixtos para IMACEC y TPM, ya que determinadas especificaciones de la prueba rechazan la hipótesis nula de raíz unitaria mientras que otras especificaciones no permiten el rechazo de dicha hipótesis. Por el contrario, para las series de Ventas e Inflación los resultados no permiten rechazar la hipótesis de raíz unitaria, en todas las especificaciones propuestas.

La estimación de los modelos de tipo ARIMA se realiza en diferencias, para asegurar la estacionariedad de las series. Se especifican hasta 24 modelos distintos, según el número de rezagos permitidos (desde 1 hasta 6) y de acuerdo a la inclusión del componente de medias móviles (desde 1 hasta 3 rezagos). El número de rezagos máximos es 6 para el componente autorregresivo y 3 para el componente de medias móviles; esto se justifica porque, según el principio de parsimonia de Box y Jenkins (1970), el número de parámetros en un modelo debería lo más reducido posible. De esta manera, en la ecuación (3), p tomaría valores de 1 a 6 y q de 1 a 3.

Para los modelos de tipo ARLI se sigue la recomendación de Demers (2005) de incluir hasta un máximo de 6 rezagos por cada variable. De esta manera, en la ecuación (4), p y q tomarían valores de 1 a 6.

Con respecto a los modelos VAR, el número máximo de rezagos permitidos se decide en base a los grados de libertad mínimos para poder realizar la estimación. Por esta razón, se estima que en la ecuación (5), p debiera tomar valores entre 1 y 12, lo cual permitiría capturar una buena parte de la dinámica de las series. Adicionalmente, se considera la posibilidad de incluir un vector de corrección del error como en la ecuación (6), con el objetivo de integrar en la estimación las potenciales relaciones de largo plazo entre las variables. No obstante, dado que tanto las pruebas de raíces unitarias como las de cointegración entregan resultados mixtos, se prefiere no hacer uso de dicha especificación. En general, estas pruebas suelen presentar problemas cuando las raíces son próximas a uno, en vez de ser exactamente uno según la hipótesis nula del contraste (Phillips, 1988), y por tanto es común que detecten relaciones de cointegración cuando realmente no existen (De Boef y Granato, 1999; Hjalmarsson y Österholm, 2007).

Para la estimación de los modelos BVAR se consideran las *priors* de Minnesota y Sims-Zha. Para esta última, se asume una distribución Normal-Wishart, tal como se describe en el apartado anterior. En ambos casos se considera el mismo número de rezagos: 3, 6, 12, 18 y 24. En la literatura sobre modelos BVAR no existe un consenso sobre cómo seleccionar el número óptimo de rezagos; por lo general se recomienda incluir tantos como sea posible, teniendo en cuenta que esto no afecta a los grados de libertad tal como ocurre en la inferencia clásica. Algunos autores, como Todd (1988a), sugieren incluir como mínimo tantos rezagos como para considerar un año entero (doce rezagos en el caso de datos con frecuencia mensual).

Cabe destacar que, contrario al enfoque clásico de inferencia, la presencia de tendencias en las variables incluidas en el modelo no genera problemas en el enfoque bayesiano (Canova 1999). Esto se debe a que la función de verosimilitud conserva la forma gaussiana a pesar de que existan raíces unitarias, por lo tanto la estimación del modelo puede hacerse en niveles (Sims, 1988).

Para el modelo BVAR con *prior* de Minnesota se consideran seis especificaciones diferentes para los hiper-parámetros, de acuerdo a la Tabla 1. En general, los hiper-parámetros a_1, a_2 y a_3 controlan el ajuste de las creencias *a priori* sobre los coeficientes, de tal manera que cuanto más próximos a cero menor es la incertidumbre. Las especificaciones propuestas para los hiper-parámetros otorgan distinto peso al conocimiento previo sobre los coeficientes: el modelo A sería el menos incierto, mientras que los modelos D, E y F los que cuentan con mayor incertidumbre.

Para el modelo BVAR con *prior* Sims-Zha también se consideran seis especificaciones para los hiper-parámetros, definidos en la Tabla 2. Los hiper-parámetros de mayor interés son λ_1 y λ_2, que controlan el ajuste de las creencias *a priori* sobre los coeficientes, y λ_3 que determina en qué medida los rezagos más antiguos pierden poder explicativo. En este caso, los modelo A y E serían los de mayor incertidumbre, con la diferencia de que en E los rezagos pierden importancia más rápidamente. Los modelos B y C serían moderadamente inciertos, mientras que D y F serían los de menor incertidumbre. Los valores de los hiper-parámetros para interceptos, exógenas, raíz unitaria y cointegración se dejan inalterados con respecto a los valores estándar propuestos por sus autores.

Tabla 1. Definición de los hiper-parámetros de la *prior* de Minnesota

	Definición	Modelo A	Modelo B	Modelo C	Modelo D	Modelo E	Modelo F
a_1	Peso otorgado a la variable dependiente	0,1	0,2	0,3	0,4	0,5	0,2
a_2	Peso otorgado al resto de variables	0,1	0,2	0,3	0,4	0,2	0,5
a_3	Escala sobre los interceptos	1	1	1	1	1	1
B_0	Media de los coeficientes	0	0	0	0	0	0

Fuente: Elaboración propia

Tabla 2. Definición de los hiper-parámetros de la *prior* de Sims-Zha

	Definición	Modelo A	Modelo B	Modelo C	Modelo D	Modelo E	Modelo F
λ_0	Ajuste general de la *prior*	1	0,8	1	0,7	1	0,8
λ_1	Ajuste sobre el componente AR(1)	1	1	0,8	0,7	1	0,8
λ_3	Contracción de los rezagos	1	1	1	1	2	4
λ_4	Ajuste sobre interceptos	1	1	1	1	1	1
λ_5	Ajuste sobre exógenas	0,07	0,07	0,07	0,1	0,07	0,07
μ_5	Ajuste sobre raíz unitaria	0,1	0,1	0,1	0,1	0,1	0,1
μ_6	Ajuste sobre cointegración	0,1	0,1	0,1	0,1	0,1	0,1

Fuente: Elaboración propia

4. Evaluación de las proyecciones

Tradicionalmente, la evaluación de las predicciones se ha realizado a partir de la especificación de una función de pérdida que mide cuánto se aleja el pronóstico del valor real. Existen varios indicadores que pueden utilizarse como función de pérdida, siendo el error cuadrático medio (ECM) uno de los más utilizados en econometría. El procedimiento común es designar como ganador al modelo que entrega el menor error cuadrático medio, pues se entiende que sus pronósticos son los más cercanos al valor real de la variable.

En las tablas 5, 6, 7, 8 y 9 del Anexo se presentan los resultados del error cuadrático medio de cada uno de los modelos estimados. Se puede apreciar que en las proyecciones a los distintos plazos, los dos modelos BVAR presentan errores, en la mayoría de los casos, menores que las alternativas de estimación clásicas.

El modelo BVAR con *prior* de Minnesota entrega las mejores predicciones, en los tres plazos considerados, cuando el número de rezagos es igual a seis. Solo en una ocasión los modelos con 12 y 18 rezagos ofrecen menor ECM, mientras que los modelos con 24 rezagos en ninguna ocasión minimizan el ECM. Con respecto al tipo de ajuste de los hiper-parámetros, la predicción a corto plazo (tres meses adelante) favorece a las especificaciones más ajustadas (menor incertidumbre), mientras que las predicciones a mayor plazo (seis y doce meses adelante) son mejores en los modelos más flexibles (mayor incertidumbre) y que, además, otorgan mayor influencia a los rezagos de la propia variable.

El modelo BVAR con *prior* de Sims-Zha entrega las mejores predicciones, mayoritariamente, cuando el número de rezagos es igual a doce. En pocas ocasiones los modelos con 6 y 18 rezagos ofrecen menor ECM, mientras que únicamente una vez el modelo con 24 rezagos minimiza el ECM. Con respecto al tipo de ajuste de los hiper-parámetros, la predicción a corto y medio plazo (tres y seis meses adelante) favorecen a las especificaciones más flexibles (mayor incertidumbre), mientras que las predicciones a mayor plazo (doce meses adelante) son más precisas en los modelos más rígidos (menor incertidumbre) y con una pérdida de importancia de los rezagos más temprana.

Si se comparan los cinco tipos de modelos estimados, se observa que, para el horizonte de predicción de 12 meses, el modelo BVAR con *prior* Sims-Zha es el que ofrece el menor error cuadrático medio (concretamente, la especificación F que impone un decaimiento más rápido en la importancia de los rezagos), seguido por el modelo ARLI con un ECM 6% superior. Debe mencionarse que el modelo BVAR con *prior* Minnesota no consigue mejorar la predicción que entrega el VAR clásico, lo cual puede deberse a una especificación poco precisa de los hiper-parámetros.

Para el horizonte de predicción de seis meses, los modelos BVAR con *prior* Minnesota (especificación E, más flexible y con mayor peso de los rezagos propios) y VAR clásico ofrecen pronósticos igual de precisos. Le siguen el BVAR con *prior* Sims-Zha (especificación E, más flexible y con decaimiento más rápido en los rezagos) y el modelo ARLI, que presentan un ECM 11% superior.

Por último, para el horizonte de predicción de tres meses, los modelos clásicos VAR y ARLI superan a los bayesianos, aunque la diferencia no es tan relevante (ECM 4% superior).

Si bien la comparación del error cuadrático medio entregado por cada modelo es un criterio generalmente aceptado para la evaluación de predicciones, puede ocurrir que la diferencia en este indicador no sea estadísticamente significativa, con lo cual no sería válido seleccionar un modelo sobre otro en base a un error cuadrático medio menor. Por esta razón, se utiliza el test de Diebold y Mariano (1995) para comparar estadísticamente las diferencias en la habilidad predictiva entre los modelos propuestos. Sean $\{e_t^i\}_{t=1}^T$ y $\{e_t^j\}_{t=1}^T$ los errores de predicción de los modelos i y j, el test estadístico se define como $S = d/\sigma_d$, donde d es el promedio del diferencial de la función de pérdida, $\{d_t\}_{t=1}^T$ se obtiene mediante $d_t = (e_t^i)^2 - (e_t^j)^2$, y σ_d es el error estándar de d. El estadístico S se distribuye asintóticamente como una variable aleatoria normal estándar bajo la hipótesis nula de igual capacidad predictiva (d=0). Valores positivos del estadístico sugieren superioridad del modelo j frente al i en términos de la predicción fuera de muestra.

Los resultados del test de Diebold y Mariano se muestran en la tabla 10 del Anexo. El test se aplicó únicamente a la especificación de cada modelo que, para cada horizonte de predicción, entregó el menor error cuadrático medio. De esta manera, para cada horizonte se comparan los cinco modelos seleccionados según este criterio, resultando un total de diez comparaciones. Se observa que, en general, los modelos bayesianos presentan mejoras estadísticamente significativas en sus predicciones fuera de muestra al compararlos con los modelos clásicos, lo cual es más evidente en los horizontes de predicción de más largo plazo.

En el horizonte de predicción de tres meses, únicamente el modelo bayesiano con *prior* de Minnesota presenta una diferencia significativa frente a las alternativas clásicas. En el horizonte de seis meses, en cambio, tanto la *prior* de Minnesota como la de Sims-Zha entregan predicciones más precisas; cabe destacar que la comparación entre ambas *prior* favorece a la de Minnesota de acuerdo al test de Diebold y Mariano. Por último, en el horizonte de doce meses es la *prior* de Sims-Zha la más destacada, pues ofrece pronósticos más certeros que los modelos alternativos ARLI y VAR, aunque no logra superar la predicción más precisa entregada por la *prior* de Minnesota.

Estos resultados son coherentes con la evidencia existente en la literatura acerca de la mejora en la precisión de las proyecciones que puede obtenerse al aplicar restricciones de tipo bayesiano en los vectores autorregresivos. En la mayoría de las aplicaciones revisadas, los modelos BVAR consiguen superar a los modelos clásicos, ya sean univariantes o vectores autorregresivos. Con respecto a las diferencias entre las dos *prior* empleadas, si el objetivo es la predicción a largo plazo, la especificación de Sims-Zha parece entregar mejores resultados, reduciendo el ECM hasta en 22% en comparación con la *prior* de Minnesota. Este resultado está en línea con lo reportado por Robertson y Tallman (1999) y Sevinç y Ergün (2009), quienes encuentran que la *prior* de Sims-Zha es más precisa en sus pronósticos.

5. Conclusiones

Con la finalidad de mejorar la predicción de venta de viviendas, se planteó la evaluación de cinco tipos de modelos econométricos ampliamente usados en la literatura internacional. Desde el más sencillo, el puramente autorregresivo, hasta el más complejo, que incorpora técnicas bayesianas en su estimación, se consideran más de 150 especificaciones distintas, las cuales se comparan mediante el error cuadrático medio y el test Diebold-Mariano para verificar cuál de ellas es más precisa en sus pronósticos.

Los resultados obtenidos están en línea con la evidencia existente en la literatura, ya que muestran que es posible mejorar la precisión de las predicciones al imponer restricciones de tipo bayesiano en un vector autorregresivo. De esta manera, los modelos BVAR estimados ofrecen mejores resultados que los modelos clásicos, por lo menos en horizontes de predicción de medio y largo plazo (6 y 12 meses, respectivamente). Esta mejora en la precisión se obtiene, generalmente, al emplear hiper-parámetros más flexibles, de tal manera que los modelos presenten mayor grado de incertidumbre sobre sus coeficientes.

Este resultado es positivo por dos razones. En primer lugar, porque los vectores autorregresivos de tipo bayesiano, como los empleados en este ejercicio, suponen una mejora casi generalizada con respecto a otros modelos comúnmente empleados en econometría, y por tanto esta metodología puede ser aplicada para pronosticar cualquier serie económica relevante. En segundo lugar, los modelos BVAR especificados en este documento muestran mayor precisión al compararlos con el modelo empleado hasta el momento (ARLI), lo cual implica que sería posible publicar pronósticos más precisos.

No obstante, existe margen para enriquecer aún más los modelos BVAR propuestos. En este ejercicio se define un conjunto de hiper-parámetros estándar, lo cual puede ser mejorado utilizando técnicas de calibración más avanzadas, que no consideren únicamente la minimización del error cuadrático medio de las predicciones. Más aún, los modelos utilizados podrían ampliarse mediante la inclusión de variables adicionales, de tal manera que se defina una ecuación de demanda por vivienda más completa que la empleada en este ejercicio. También podría hacerse uso de otros métodos avanzados, como modelos de combinación de predicciones. Además, podría explorarse la capacidad predictiva de otros modelos, por ejemplo los denominados FAVAR (Factor-Augmented Vector Autoregressive) desarrollados por Bernanke *et al.* (2005).

Referencias

Banbura, M., Giannone, D., & Reichlin, L. (2010). Large Bayesian Vector Auto Regressions. Journal of Applied Econometrics, 25(1), 71–92.

Bernanke, B., Boivin, J., & Eliasz, P. (2005). Measuring monetary policy: a factor augmented autoregressive (FAVAR) approach. Quarterly Journal of Economics, 120, 387–442.

Box, G., & Jenkins, G. (1970). Time Series Analysis: Forecasting and Control. Holden Day, San Francisco.

Canova, F. (1999). Vector Autoregressive Models: Specification, Estimation, Inference, and Forecasting. Handbook of Applied Econometrics. Volume I: Macroeconomics (M. Hashem Pesaran and Michael R. Wickens.). Blackwell Publishing.

Canova, F. (2007). Bayesian VARs. Methods for Applied Macroeconomic Research (pp. 351–394). Princeton University Press.

Ciccarelli, M., & Rebucci, A. (2003). Bayesian VARs: A Survey of the Recent Literature with an Application to the European Monetary System. Working Paper WP/03/102, International Monetary Fund.

Crone, T. M., & McLaughlin, M. P. (1999). A Bayesian VAR Forecasting Model for the Philadelphia Metropolitan Area. Working Paper No. 99-7, Federal Reserve Bank of Philadelphia.

De Boef, S., & Granato, J. (1999). Testing for Cointegrating Relationships with Near-Integrated Data. Political Analysis, 8(1), 99–117.

Demers, F. (2005). Modelling and Forecasting Housing Investment: The Case of Canada. Working Paper 2005-41, Bank of Canada.

Diebold, F. X. (2007). Forecasting with Regression Models. Elements of forecasting (4th ed., pp. 219–256). Thomson/South-Western.

Diebold, F. X., & Mariano, R. (1995). Comparing Predictive Accuracy. Journal of Business & Economic Statistics, 13(3), 253–263.

Doan, T., Litterman, R., & Sims, C. A. (1983). Forecasting and Conditional Projection Using Realistic Prior Distributions. NBER Working Paper No. 1202.

Dua, P., Miller, S. M., & Smyth, D. J. (1996). Using Leading Indicators to Forecast US Home Sales in a Bayesian VAR Framework. Working papers 1996-08, University of Connecticut, Department of Economics.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. Econometrica, Econometric Society, 55(2), 251–76.

Garcia-Ferrer, A., Highfield, R.A., Palm, F., & Zellner, A. (1987). Macroeconomic Forecasting Using Pooled International Data. Journal of Business & Economic Statistics, 5(1), 53–67.

Giannone, D., Lenza, M., & Primiceri, G. E. (2012). Prior Selection for Vector Autoregressions. NBER Working Paper No. 18467.

González, W. (2012). Un Gran VAR Bayesiano para la Economía Chilena. Documento de Trabajo No 653, Banco Central de Chile.

Gupta, R., Jurgilas, M., Kabundi, A., & Miller, S. M. (2009). Monetary Policy and Housing Sector Dynamics in a Large-Scale Bayesian Vector Autoregressive Model. Working Paper 2009-19, Department of Economics Working Paper Series, University of Connecticut.

Gupta, R., & Sichei, M. M. (2006). A BVAR Model for the South African Economy. Working Papers 200612, University of Pretoria, Department of Economics.

Hjalmarsson, E., & Österholm, P. (2007). Testing for Cointegration Using the Johansen Methodology when Variables are Near-Integrated. IMF Working Paper WP/07/141.

Jaramillo, P. (2009). Estimación de VAR bayesianos para la economía chilena. Revista de Análisis Económico, 24(1), 101–126.

Johnston, J., & DiNardo, J. (1997). Univariate Time Series Modeling. Econometric Methods (4th ed., pp. 204–243). McGraw-Hill.

Joiner, A. (2001). Monetary policy effects in an Australian Bayesian VAR model. Working Paper, Departament of Econometrics and Business Statistics, Monash University.

Kadiyala, K. R., & Karlsson, S. (1997). Numerical Methods For Estimation and Inference in Bayesian VAR Models. Journal of Applied Econometrics, 12, 99–132.

Keller, E. (2007). Classical and Bayesian Methods for the VAR Analysis: International Comparisons. Rivista di Politica Economica, 97(6), 149–202.

Kenny, G., Meyler, A., & Quinn, T. (1998). Bayesian Var Models for Forecasting Irish Inflation. Research Technical Papers 4/RT/98, Central Bank of Ireland.

Koop, G. (2003). Bayesian Econometrics. John Wiley & Sons Ltd.

Koop, G., & Korobilis, D. (2009). Bayesian Multivariate Time Series Methods for Empirical Macroeconomics. Working Paper Series 47_09, The Rimini Centre for Economic Analysis.

Leamer, E. E. (2007). Housing is the business cycle. NBER Working Paper No. 13428.

Litterman, R. (1979). Techniques of Forecasting Using Vector Autoregressions. Working Paper No. 15, Federal Reserve Bank of Minneapolis.

Litterman, R. (1986). Forecasting With Bayesian Vector Autoregressions - Five Years of Experience. Journal of Business & Economic Statistics, 4(1), 25–38.

Llosa, G., Tuesta, V., & Vega, M. (2006). Un modelo de proyección BVAR para la inflación peruana. Estudios Económicos No 13, Banco Central de Reserva del Perú.

Phillips, P. C. B. (1988). Regression Theory for Near-Integrated Time Series. Econometrica, 56(5), 1021–1043.

Robertson, J. C., & Tallman, E. W. (1999). Vector Autoregressions: Forecasting and Reality. Economic Review, First Quarter 1999, Federal Reserve of Atlanta.

Sevinç, V., & Ergün, G. (2009). Usage of different prior distributions in bayesian vector autoregressive models. Hacettepe Journal of Mathematics and Statistics, 38(1), 85–93.

Sims, C. A. (1980). Macroeconomics and Reality. Econometrica, 48(1), 1–47.

Sims, C. A. (1988). Bayesian skepticism on unit root econometrics. Discussion Paper 3, Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis.

Sims, C. A., & Zha, T. (1996). Bayesian Methods for Dynamic Multivariate Models. Working Paper 96-13, Federal Reserve Bank of Atlanta.

Todd, R. M. (1988a). Implementing Bayesian Vector Autoregressions. Working Paper 384, Federal Reserve Bank of Minneapolis.

Todd, R. M., & Morande, F. G. (1988a). A BVAR Forecasting Model for the Chilean Economy. Revista de Análisis Económico, 3(2), 45–78.

Zellner, A., & Hong, C. (1989). Forecasting international growth rates using Bayesian shrinkage and other procedures. Journal of Econometrics, 40(1), 183–202.

Zellner, Arnold, & Tobias, J. (2000). A Note on Aggregation, Disaggregation and Forecasting Performance. Staff General Research Papers, Iowa State University, Department of Economics.

Anexos

Tabla 3. Prueba de raíz unitaria (Test Dickey-Fuller Aumentado)

	Con constante		Constante y pendiente		Sin constante ni pendiente	
	Estadístico	P-valor	Estadístico	P-valor	Estadístico	P-valor
Ventas	-1,793	0,38	-2,254	0,46	0,101	0,71
Imacec	-3,043	0,03	-3,127	0,10	-1,675	0,09
TPM	-2,580	0,10	-2,966	0,14	-1,143	0,23
Inflación	-2,387	0,15	-2,330	0,42	-1,925	0,05

Fuente: Elaboración propia

Tabla 4. Prueba de raíz unitaria (Dickey-Fuller GLS)

	Con constante				Constante y pendiente			
		Valores críticos				Valores críticos		
	Estadístico	1%	5%	10%	Estadístico	1%	5%	10%
Ventas	-1,54	-2,58	-1,94	-1,62	-2,19	-3,46	-2,93	-2,64
Imacec	-3,05	-2,58	-1,94	-1,62	-3,09	-3,46	-2,93	-2,64
TPM	-2,22	-2,58	-1,94	-1,62	-2,98	-3,46	-2,93	-2,64
Inflación	-0,40	-2,58	-1,94	-1,62	-1,47	-3,46	-2,93	-2,64

Fuente: Elaboración propia

Tabla 5. Modelos ARIMA: Error Cuadrático Medio

ECM	Horizonte (Muestra Ene94-Jun11)		
Modelos	h = 3	h = 6	h = 12
AR(1)	0.081	0.068	0.158
AR(2)	0.081	0.069	0.162
AR(3)	0.082	0.067	0.154
AR(4)	0.082	0.069	0.157
AR(5)	0.082	0.069	0.157
AR(6)	0.082	0.070	0.156
ARMA(1,1)	0.081	0.068	0.159
ARMA(1,2)	0.079	0.067	0.159
ARMA(1,3)	0.080	0.082	0.191
ARMA(2,1)	0.082	0.069	0.161
ARMA(2,2)	0.090	0.071	0.155
ARMA(2,3)	0.097	0.084	0.182
ARMA(3,1)	0.079	0.083	0.195
ARMA(3,2)	0.079	0.082	0.192
ARMA(3,3)	0.087	0.074	0.167
ARMA(4,1)	0.085	0.067	0.150
ARMA(4,2)	0.084	0.072	0.164
ARMA(4,3)	0.097	0.083	0.172
ARMA(5,1)	0.081	0.070	0.159
ARMA(5,2)	0.084	0.070	0.160
ARMA(5,3)	0.099	0.086	0.177
ARMA(6,1)	0.080	0.067	0.155
ARMA(6,2)	0.082	0.072	0.168
ARMA(6,3)	0.098	0.087	0.179

Fuente: Elaboración propia

Tabla 6. Modelos ARLI: Error Cuadrático Medio

ECM	Horizonte (Muestra Ene94- Jun11)		
Modelos	h = 3	h = 6	h = 12
ARLI(1,1)	0.078	0.076	0.179
ARLI(1,2)	0.089	0.068	0.141
ARLI(1,3)	0.089	0.068	0.130
ARLI(1,4)	0.087	0.073	0.109
ARLI(1,5)	0.077	0.077	0.115
ARLI(1,6)	0.077	0.079	0.118
ARLI(2,1)	0.078	0.073	0.170
ARLI(2,2)	0.091	0.069	0.137
ARLI(2,3)	0.090	0.068	0.129
ARLI(2,4)	0.087	0.073	0.107
ARLI(2,5)	0.076	0.078	0.116
ARLI(2,6)	0.081	0.081	0.116
ARLI(3,1)	0.078	0.072	0.167
ARLI(3,2)	0.090	0.069	0.138
ARLI(3,3)	0.088	0.067	0.129
ARLI(3,4)	0.083	0.069	0.108
ARLI(3,5)	0.072	0.075	0.114
ARLI(3,6)	0.069	0.071	0.116
ARLI(4,1)	0.078	0.080	0.178
ARLI(4,2)	0.082	0.066	0.144
ARLI(4,3)	0.082	0.067	0.135
ARLI(4,4)	0.073	0.057	0.108
ARLI(4,5)	0.059	0.059	0.115
ARLI(4,6)	0.061	0.061	0.119
ARLI(5,1)	0.076	0.074	0.170
ARLI(5,2)	0.084	0.066	0.139
ARLI(5,3)	0.084	0.067	0.130
ARLI(5,4)	0.072	0.056	0.111
ARLI(5,5)	0.064	0.064	0.118
ARLI(5,6)	0.063	0.064	0.119
ARLI(6,1)	0.077	0.080	0.177
ARLI(6,2)	0.082	0.070	0.147
ARLI(6,3)	0.082	0.070	0.138
ARLI(6,4)	0.070	0.055	0.114
ARLI(6,5)	0.063	0.060	0.121
ARLI(6,6)	0.059	0.057	0.124

Fuente: Elaboración propia

Tabla 7. Modelos VAR: Error Cuadrático Medio

ECM	Horizonte (Muestra Ene94- Jun11)		
Modelos	h = 3	h = 6	h = 12
VAR(1)	0.072	0.078	0.188
VAR(2)	0.076	0.067	0.170
VAR(3)	0.080	0.058	0.143
VAR(4)	0.092	0.070	0.129
VAR(5)	0.078	0.067	0.122
VAR(6)	0.084	0.065	0.116
VAR(7)	0.064	0.050	0.116
VAR(8)	0.066	0.056	0.126
VAR(9)	0.068	0.063	0.141
VAR(10)	0.054	0.068	0.163
VAR(11)	0.057	0.064	0.166
VAR(12)	0.099	0.089	0.155

Fuente: Elaboración propia

Tabla 8. Modelos BVAR (Minnesota) Error Cuadrático Medio

ECM	Horizonte (Muestra Ene94-Jun11)		
Modelos	h = 3	h = 6	h = 12
A(3)	0.096	0.139	0.303
B(3)	0.083	0.078	0.176
C(3)	0.083	0.072	0.180
D(3)	0.080	0.077	0.178
E(3)	0.078	0.074	0.182
F(3)	0.078	0.073	0.174
A(6)	0.061	0.069	0.171
B(6)	0.062	0.054	0.155
C(6)	0.069	0.055	0.146
D(6)	0.073	0.055	0.130
E(6)	0.074	0.050	0.143
F(6)	0.070	0.057	0.149
A(12)	0.084	0.100	0.188
B(12)	0.072	0.090	0.170
C(12)	0.072	0.077	0.157
D(12)	0.068	0.074	0.153
E(12)	0.068	0.073	0.143
F(12)	0.075	0.082	0.161
A(18)	0.108	0.112	0.194
B(18)	0.094	0.095	0.175
C(18)	0.080	0.077	0.160
D(18)	0.068	0.071	0.150
E(18)	0.075	0.072	0.158
F(18)	0.084	0.083	0.163
A(24)	0.089	0.108	0.206
B(24)	0.080	0.092	0.180
C(24)	0.076	0.081	0.162
D(24)	0.074	0.072	0.151
E(24)	0.076	0.082	0.163
F(24)	0.076	0.080	0.163

Fuente: Elaboración propia

Tabla 9. Modelos BVAR (Sims-Zha) Error Cuadrático Medio

ECM	Horizonte (Muestra Ene94- Jun11)		
Modelos	h = 3	h = 6	h = 12
A(3)	0.083	0.073	0.176
B(3)	0.081	0.069	0.164
C(3)	0.080	0.072	0.166
D(3)	0.082	0.073	0.160
E(3)	0.082	0.074	0.175
F(3)	0.079	0.071	0.171
A(6)	0.081	0.067	0.121
B(6)	0.086	0.063	0.115
C(6)	0.078	0.063	0.118
D(6)	0.081	0.069	0.114
E(6)	0.082	0.059	0.119
F(6)	0.086	0.059	0.132
A(12)	0.066	0.058	0.133
B(12)	0.070	0.057	0.130
C(12)	0.069	0.060	0.126
D(12)	0.070	0.057	0.121
E(12)	0.073	0.055	0.115
F(12)	0.085	0.065	0.117
A(18)	0.151	0.199	0.204
B(18)	0.150	0.192	0.191
C(18)	0.150	0.186	0.192
D(18)	0.130	0.163	0.159
E(18)	0.090	0.088	0.115
F(18)	0.080	0.063	0.106
A(24)	0.171	0.233	0.258
B(24)	0.170	0.223	0.244
C(24)	0.166	0.221	0.244
D(24)	0.150	0.195	0.204
E(24)	0.100	0.118	0.129
F(24)	0.074	0.061	0.101

Fuente: Elaboración propia

Tabla 10. Test Diebold-Mariano

h=3		ARLI(6,6)	VAR(10)	Minnesota A(6)	Sims-Zha A(12)
	ARMA(3,1)	2.316**	8.553***	6.967***	1,28
	ARLI(6,6)		0,3661	-0,2574	-1,289
	VAR(10)			-1,598	-0,8772
	Minnesota A(6)				-0,5281
h=6		ARLI(6,4)	VAR(7)	Minnesota E(6)	Sims-Zha E(12)
	AR(3)	3.491***	6.711***	7.296***	5.693***
	ARLI(6,4)		1,258	2.632***	-0,3028
	VAR(7)			-0,05715	-2.667***
	Minnesota E(6)				-5.848***
h=12		ARLI(2,4)	VAR(6)	Minnesota D(6)	Sims-Zha F(24)
	ARMA(4,1)	1,277	1,402	1.814*	1,478
	ARLI(2,4)		-0,8869	-0,9619	2.718***
	VAR(6)			-1,004	1.703**
	Minnesota D(6)				1,267

(***) Significativo al 1%, (**) Significativo al 5%, (*) Significativo al 10%

Fuente: Elaboración propia