

PREDICTORES DEL IMACEC

Juan Facundo Piguillem
Economista
Gerencia de Estudios
Cámara Chilena de la Construcción

Resumen

En este trabajo se estiman y evalúan las técnicas habituales para proyectar series de tiempo, aplicándolas exclusivamente al Estimador Mensual de Actividad Económica (IMACEC). Con este fin, se estiman tres clases de modelos de proyección, un indicador líder, un coincidente y un modelo VAR. La superioridad de uno sobre otro no aparece clara, lo que dependerá de la calidad de información de la que se disponga y del retraso con la que aparece.

La publicación de los Documentos de Trabajo no está sujeta a la aprobación previa de la Mesa Directiva de la Cámara Chilena de la Construcción. Tanto el contenido de los Documentos de Trabajo, como también el análisis y conclusiones que de ellos se deriven, son exclusiva responsabilidad de su(s) autor(es) y no reflejan necesariamente la opinión de la Cámara Chilena de la Construcción o sus directivos.

INDICE

I. Introducción	Página 2
II. Los modelos	
II.1. Modelos Univariados	Página 2
II.2. Modelos de ecuaciones simultaneas	Página 8
III. Elección del Predictor	Página 12
IV. Resumen y Conclusiones	Página 13
Bibliografía	Página 14

I. Introducción

Desde los inicios de la ciencia económica uno de los aspectos que más ha concitado la atención en el área ha sido la predicción o proyección de las variables económicas de mayor relevancia, tales como la inflación, el nivel de actividad económica, la tasa de interés, etc. Dentro de éstas, quizás la de mayor importancia sea el nivel de actividad económica. Para proyectarla, en general, se recurre a indicadores ya confeccionados por organismos oficiales. En Chile existen tres medidas del nivel de actividad de acuerdo al plazo que abarcan: el Producto Bruto Interno (PIB) anual, el PIB trimestral y el Indicador Mensual de Actividad Económica (IMACEC). La precisión con la que cada uno de éstos miden el nivel de actividad efectivo se encuentra inversamente relacionado con el período que comprende la medición. Así, el PIB anual da la medida más exacta, en tanto que el IMACEC, de medición mensual, es el que puede derivar en mayor cantidad de errores.

En este documento se aplicarán diferentes metodologías para proyectar el IMACEC, que si bien es un indicador menos preciso que los anteriores, permite monitorear el estado de la economía con mayor frecuencia que el PIB trimestral y anual. Sin embargo, dado que el IMACEC se conoce con un mes de retraso, las "proyecciones" se realizarán no sólo para sus valores futuros sino que también para valores presentes y pasados. Claramente las proyecciones sobre estos últimos, llamados indicadores coincidentes, serán más precisas que las efectuadas para los valores futuros, debido a que en el primer caso es posible utilizar información contemporánea respecto a otras variables que se conocen con anticipación al IMACEC, y que finalmente lo componen.

II. Los modelos

II.1. Modelos Univariados

Existen dos clases principales de modelos de predicción de variables en el tiempo: modelos uniecuacionales (modelos ARIMA, líder, coincidente, etc.), y modelos multiecuacionales o de sistemas, también conocidos como VAR (*Vector Autoregression Models*). A pesar de la mayor complejidad de los VAR, la superioridad de éstos sobre los modelos univariados no es clara. Si bien es cierto que los modelos VAR son más ricos, en cuanto a la dinámica que generan, y poseen en general mejores propiedades estadísticas, ya que permiten la retroalimentación entre las variables intervinientes, para algunas series de tiempo se encuentra que sus valores futuros son mejor predichos utilizando solamente la información que la misma variable ha entregado en el pasado. Así, como una primera aproximación, y por ser el modelo más sencillo, para predecir el IMACEC

se ajusta un modelo ARIMA, el cual utiliza como únicos argumentos los valores que haya arrojado la misma variable en el pasado. Matemáticamente la ecuación de proyección sería:

$$Y_t = \alpha + \sum_{i=1}^p \rho_i Y_{t-i} + \sum_{j=1}^q \phi_j \varepsilon_{t-j} + \varepsilon_t \quad (1)$$

Donde Y_t es la variable sobre la que se desea realizar la predicción, en este caso el IMACEC, y ε_t es un error aleatorio que se distribuye normal, con media cero y varianza σ^2 ($\varepsilon_t \sim N(0, \sigma^2)$). Además se debe satisfacer que el error sea un *Ruido Blanco*, es decir, que no este correlacionado con Y_t , ni con sus valores pasados. Otro supuesto importante es que la variable sea estacionaria, lo que implica una restricción sobre los parámetros asociados a los valores rezagados de Y_t , específicamente que $-1 < \rho_i < 1$. Si esta restricción no se cumple se dice que variable sigue un *Camino Aleatorio*¹.

Cuando este último supuesto no se cumple se dice que la serie es integrada o no estacionaria, en cuyo caso la práctica habitual es tomar las diferencias temporales de la variable hasta que el supuesto se cumpla. En el caso del IMACEC, tomando los valores de la variable, tanto en niveles como en el logaritmo del nivel, no se puede rechazar la hipótesis nula de raíz unitaria. A tal fin se trabaja con la diferencia en doce meses del logaritmo de Y_t , variable que de ahora en más se identificará como C12IMA. Como se observa en la tabla inferior, aplicando la prueba de Philip-Perron para la detección de raíces unitarias, suponiendo la existencia de intercepto y que la variable no posee tendencia, se rechaza la hipótesis nula de raíz unitaria con un 97% de confianza.

Tabla N° 1
Test de Raíz Unitaria sobre C12IMA

	5% Critical Value	-3.4497
	10% Critical Value	-3.1499
*MacKinnon critical values for rejection of hypothesis of a unit root.		

El modelo presentado en la ecuación (1) se denomina ARMA(p,q)², donde la primera parte de la expresión (AR), o Auto Regresivo, hace referencia a los términos rezagados de la variable consigo misma, representado por los ρ_i de la ecuación (1), en tanto que MA, los términos de Media Móvil, hacen referencia a los rezagos del error, reflejados en los ϕ_j de la misma ecuación. Para la estimación de

¹ Para una demostración de esta afirmación ver Green (1998)

² Siempre que se cumpla el supuesto de ausencia de raíz unitaria, caso contrario sería una ARIMA.

los parámetros de (1), se requiere en primer lugar encontrar el número de rezagos a incluir en la estimación, i.e., los valores de j e i . Esta tarea se ha realizado utilizando los criterios de Akaike y Schwarz.

Utilizando datos mensuales que abarcan desde enero de 1992 hasta junio de 2002 se estimó la ecuación (1), la que se presenta en la siguiente tabla.

Tabla N° 2
Modelo ARMA

Variable Dependiente: C12IMA ³				
Muestra(ajustada): 1993:02 2002:06				
Observaciones Incluidas : 113				
Backcast: 1992:05 1993:01				
Variable	Coefficiente	Std. Error	t-Statistic	Prob.
C	0.046584	0.017540	2.655909	0.0091
AR(1)	0.859921	0.064704	13.29013	0.0000
MA(1)	-0.366731	0.088088	-4.163216	0.0001
MA(3)	0.333677	0.083837	3.980079	0.0001
MA(6)	0.180974	0.088761	2.038883	0.0439
MA(9)	0.373002	0.084632	4.407339	0.0000
R-squared	0.800838	Akaike info criterion	-5.219848	
Adjusted R-squared	0.791531	Schwarz criterion	-5.075031	
Log likelihood	300.9214	F-statistic	86.05002	
Durbin-Watson stat	2.027892	Prob(F-statistic)	0.000000	

Como se desprende de la Tabla N° 2, la variable sigue un proceso ARMA(1,9). Así, el crecimiento anual del IMACEC muestra una alta persistencia, en la ausencia de shocks externos, lo que se refleja en que cada punto porcentual de crecimiento o decrecimiento en el período anterior conlleva 0,85 puntos porcentuales en el periodo siguiente. La regresión es estadísticamente significativa ($\text{Prob}(F\text{-statistic})=0$) y los residuos son ruido blanco ($\text{corr}(\varepsilon_t, \varepsilon_{t-i})=0$, para $i=1,2,\dots,T$). Sin embargo, aunque el R^2 es elevado, todavía un 20% de la variabilidad del crecimiento del IMACEC no es explicado por factores capturados dentro del modelo.

Para subsanar esta debilidad y mejorar en lo posible las proyecciones existen dos alternativas, las que consisten básicamente en incluir en la ecuación (1) variables "exógenas", encontrándose la diferencia en el tipo y forma en la que se incorporan las variables. Así, las variables exógenas pueden ser de dos tipos: variables exógenas contemporáneas, tales como las importaciones, los índices de producción industrial y minera, etc.; y variables exógenas rezagadas, también llamadas indicadores líderes, dentro de los cuales pueden incluirse los cambios pasados en la tasa de política monetaria, en el precio del cobre, en el índice de precios de las acciones, etc.

³ La serie C12IMA es la variación porcentual del IMACEC a doce meses sobre la serie sin desestacionalizar.

Si la ecuación a estimar incluye sólo variables exógenas rezagadas, el modelo resultante es un indicador líder. De esta manera se plantean dos modelos alternativos:

$$Y_t = \alpha + \beta LX_t + \varepsilon_t \quad (2)$$

$$Y_t = \alpha + \sum_{i=1}^p \rho_i Y_{t-i} + \sum_{j=1}^q \phi_j \varepsilon_{t-j} + \beta X_t + \pi LX_t + \varepsilon_t \quad (3)$$

Donde X_t es una matriz que incluye las variables exógenas contemporáneas, LX_t representa la matriz de variables exógenas rezagadas y β es el vector de parámetros estimados.

El primer modelo, ecuación (2), se conoce como indicador líder. Trata de predecir el valor de la variable Y_t sobre la base de los movimientos o valores que mostraron ciertas variables de relevancia en el pasado. Por ejemplo, para la predicción del crecimiento del IMACEC en el momento t , serían relevantes a priori incorporar, acorde con la intuición económica, el crecimiento de la cantidad real de dinero en $t-i$, donde $i > 0$ queda por determinar, el otorgamiento de permisos de edificación también en $t-i$, la TIR de las letras hipotecarias, etc. Sin embargo, pueden existir otras variables que también sean significativas para predecir el crecimiento del IMACEC, aunque no exista una relación teórica clara, lo que no representa un problema mayor puesto que este tipo de modelos sólo busca el vínculo estadístico entre las mismas, sin ahondar en el porqué del vínculo o sus implicancias teóricas.

Por su parte, el segundo modelo, el de la ecuación (3), es una combinación de los dos primeros, incluyendo tanto características de los modelos ARMA de series de tiempo y de indicadores líderes.

La ventaja del indicador líder sobre el segundo modelo es el menor requerimiento de información actualizada. Dependiendo del largo de los rezagos incluidos en la ecuación (2), es mucho más probable que si en el momento t se desea predecir el valor del IMACEC para $t+1$, ya se disponga de la información para los componentes de LX_t . En tanto que la ventaja del segundo modelo es el aprovechamiento de la fuerte correlación contemporánea que existe entre la variación del IMACEC y algunas variables como las importaciones, el índice de producción del INE, las ventas minoristas, etc. Sin embargo, la información sobre estas últimas variables sólo se dispone con muy poca anticipación al momento en que se desea predecir. En adición, generalmente las cifras son provisionarias, lo que hace que habituales errores y futuras correcciones en las cifras publicadas inicialmente puedan trasladarse a las predicciones.

Un caso especial surge en la ecuación (3) si sólo se incluyen variables exógenas contemporáneas, en este caso en particular el modelo se denomina indicador coincidente de la actividad económica.

Es importante resaltar en este punto que, uno de los supuestos fundamentales para la correcta estimación de la ecuación (3) es que la correlación entre cada una de las variables incluidas en X_t , LX_t y el término de error sea cero, caso contrario los valores estimados para β y π serán sesgados e inconsistentes. Una condición necesaria para el cumplimiento de este supuesto es que las variables incluidas en X_t sean efectivamente exógenas. Para comprobarlo es necesario realizar un Test de Causalidad de Granger⁴ sobre cada una de ellas.

Como primer paso en la metodología para estimar el indicador líder se reunió un conjunto de 55 variables que a priori resultaban relevantes. Luego, con el objetivo de seleccionar las variables a utilizar se agrupan según los siguientes criterios⁵:

- 1) Correlación cruzada: La correlación cruzada se define como la correlación entre una variable y los rezagos de otra variable, de esta manera se tiene que,

$$\rho(i) = \frac{\text{cov}(Y_t, X_{t-i})}{\sqrt{\text{var}(Y_t)} \sqrt{\text{var}(X_{t-i})}}$$

es la correlación cruzada entre la variable Y_t , IMACEC en este caso, y la variable candidata a líder X_t con el rezago i , la que es llamada correlación cruzada de orden i . Sobre la base de este criterio, en primer lugar se busca el orden que maximiza la correlación cruzada para cada variable, el máximo $\rho(i)$ para cada X_t , y en segundo lugar, se ordenan las variables de acuerdo a las que muestran mayor correlación en valor absoluto.

- 2) Predictibilidad. Este criterio se implica elegir aquellas variables que mejor se predigan utilizando sólo la historia pasada que arroja la misma variable. Así, se ajusta un ARMA para cada candidata y se ordenan de acuerdo a la a la magnitud de su R^2 .
- 3) Causalidad de Granger: Aquí lo que se busca es detectar en que medida cada candidata causa, en el sentido de Granger, al IMACEC. Así, aquellas variables que exhiban un mayor p-value en la causalidad al IMACEC, y que no sean causadas por el mismo, son mejor valoradas.

Mediante los criterios antes mencionados se seleccionaron 20 variables con las que se calcularon tres estimadores líderes, optándose en última instancia por el que se presenta en la Tabla N° 3.

Por otra parte, para estimar el indicador coincidente, se eligieron aquellas variables que representan los sectores económicos de mayor peso en la construcción del IMACEC, y luego se construyeron diferentes modelos con combinaciones alternativas de las variables antes mencionadas, optándose por aquel modelo que minimizaba el error de predicción. La estimación resultante se encuentra en la Tabla N° 4.

⁴ En el caso de las variables incluidas en LX_t , esto no es necesario ya que por definición son exógenas, o lo que es lo mismo, están predeterminadas.

⁵ Para una explicación más extensa sobre los criterios de selección de variables "líderes" ver Bravo y Franke (2001)

Tabla Nº3 Indicador Líder

Variable Dependiente: C12IMA Método: Mínimos Cuadrados				
Variable*	Coefficient	Std. Error	t-Statistic	Prob.
C	0.042493	0.006888	6.169364	0.0000
C12MINAS(-10)	0.110620	0.030177	3.665736	0.0004
TCO_DFLAC(-2)	-0.139536	0.039561	-3.527136	0.0007
C12COBRE(-9)	0.028727	0.011351	2.530702	0.0132
STOCK(-3)	-0.001008	0.000276	-3.652484	0.0004
C12M1(-3)	0.240095	0.040400	5.943012	0.0000
CDIAL	0.007945	0.000983	8.081192	0.0000
C12TERM(-2)	0.028314	0.008113	3.499889	0.0008
C12IGPA(-8)	0.033167	0.012812	2.588817	0.0113
C12EDI(-1)	0.014086	0.005697	2.472481	0.0154
C12CAPTA(-4)	0.009016	0.005479	1.645479	0.1035
C12PRC8(-10)	0.049300	0.025541	1.930226	0.0569
MA(12)	-0.778602	0.039155	-19.88483	0.0000
MA(1)	0.272015	0.066753	4.074933	0.0001
MA(2)	0.137836	0.054418	2.532893	0.0131
R-squared	0.920746	Mean dependent var	0.048001	
Adjusted R-squared	0.907844	S.D. dependent var	0.039100	
Log likelihood	312.6166	F-statistic	71.36570	
Durbin-Watson stat	1.816258	Prob(F-statistic)	0.000000	

* La definición de cada variable se puede ver en la Tabla Nº 5

Tabla Nº 4 Indicador Coincidente

Variable Dependiente: C12IMA Muestra(ajustada): 1993:04 2002:06 Observaciones Incluidas: 111 Backcast: 1992:04 1993:03				
	Coefficient	Std. Error	z-Statistic	Prob.
C	0.038190	0.002573	14.84324	0.0000
C12_IMP	0.049778	0.003382	14.71958	0.0000
C12PRD_INE	0.268281	0.016296	16.46341	0.0000
C12MINAS	0.054727	0.008759	6.248054	0.0000
TCO_DFLAC(-2)	-0.138037	0.015255	-9.048742	0.0000
C12VTAS_INE	0.044497	0.011772	3.779881	0.0002
C12TERM	0.006566	0.003254	2.018064	0.0436
C12VTAMINOR	0.034873	0.009861	3.536478	0.0004
CDIAL	0.001504	0.000460	3.345179	0.0008
AR(1)	0.715938	0.023227	30.82329	0.0000
MA(12)	-0.462302	0.039686	-11.64884	0.0000
MA(3)	0.346574	0.033250	10.42339	0.0000
MA(2)	0.208537	0.030457	6.846946	0.0000
Ecuación de la Varianza				
C	-5.398048	1.113440	-4.848082	0.0000
[RES]/SQR[GARCH](1)	-1.567981	0.337743	-4.642531	0.0000
RES/SQR[GARCH](1)	-0.329979	0.129552	-2.547081	0.0109
EGARCH(1)	0.311505	0.133161	2.339317	0.0193
R-squared	0.941070	Akaike info criterion	-6.493735	
Adjusted R-squared	0.931040	Schwarz criterion	-6.078762	
Log likelihood	377.4023	F-statistic	93.82027	
Durbin-Watson stat	1.875292	Prob(F-statistic)	0.000000	

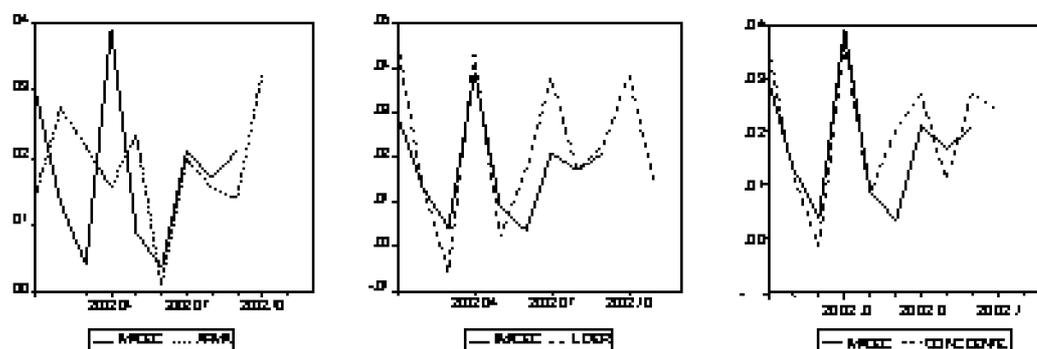
Tabla N° 5. Listado de Variables

C12_IMP :	Cambio en doce meses de las importaciones totales
C12PROD_INE :	Cambio en doce meses del índice de producción del INE
C12MINAS :	Cambio en doce meses en la producción minera
TCO_DFLAQ(-2) :	Cambio en doce meses del cociente entre el tipo de cambio observado y el índice de precios al consumidor (IPC)
C12VTAS_INE :	Cambio en doce meses del índice de ventas físicas del INE
C12TERM :	Cambio en doce meses del cociente entre los precios del cobre y del petróleo
C12VTAMINOR :	Cambio en doce meses de las ventas minoristas
CDIAL :	Diferencia de días laborales con respecto al mismo mes del año anterior
C12COBRE :	Cambio en doce meses del precio del cobre
C12PRCB :	Cambio en doce meses de la tasa del PRCB
C12CAPTA :	Cambio en doce meses de la tasa de captación entre 30 y 90 días no reajutable
C12M1 :	Cambio en doce meses del dinero real (M1)
STOCK :	Cambio con respecto a igual mes del año anterior de los meses para agotar el stock de casas y departamento
C12IGPA :	Cambio en doce meses del índice general de precios de las acciones

Algunas consideraciones merecen las estimaciones anteriores. En primer lugar, las regresiones como un todo son estadísticamente significativas y el R^2 se incrementan hasta 0,94, por lo que con los nuevos modelos se explica entre un 12% y un 14% más de la variabilidad de la tasa de crecimiento del IMACEC, respecto al modelo ARMA. Los residuos no exhiben autocorrelación, debido a la inclusión de los términos MA, y siguen una distribución normal, de acuerdo al test Jaque-Bera.

Otra de las ventajas de estos últimos sobre el modelo ARMA, es que la información que se utiliza para predecir responde a los cambios que pudiesen ocurrir en la economía previo o durante al período de predicción mejorando sustancialmente el ajuste como se aprecia en el Gráfico N° 1.

Gráfico N° 1
Proyecciones para el año 2002 con los modelos Univariados



Por último, debe mencionarse que el indicador coincidente exhibe heterocedasticidad condicional, es decir la varianza del error no es constante sino que más bien cambia dependiendo el período que se trate. El modelo para la varianza plasmado aquí es el EGARCH de primer orden. Esto supone un modelo exponencial para la varianza con memoria de un período, específicamente el patrón seguido por la varianza es:

$$\log(\sigma_t^2) = \omega + \beta \log(\sigma_{t-1}^2) + \alpha \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| + \gamma \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \quad (4)$$

Dado que γ es estadísticamente significativo el efecto es asimétrico, esto implica que en épocas de alto crecimiento la varianza del error es menor.

II.2. Modelos Dinámicos de Ecuaciones Simultáneas

Ahora bien, como se anticipó en la sección anterior, es crucial para la correcta estimación de estos modelos que la correlación entre las variables "exógenas" y la perturbación estocástica sea cero, caso contrario las estimaciones serán sesgadas e inconsistentes. El supuesto de exogeneidad de las variables "independientes" juega un rol clave en la estimación del indicador coincidente, ya que si las variables supuestamente exógenas no lo son tanto, es decir, que afectan contemporáneamente a la variable dependiente, pero al mismo tiempo se ven afectada por esta, se produce correlación entre el término de error y las variables del lado derecho de la ecuación. Esto es conocido en la literatura de ecuaciones simultaneas como sesgo de simultaneidad, y sucede siempre que a una variable endógena se la considera como exógena. Para asegurarse que el supuesto de exogeneidad se cumple se aplicó el *Test de Causalidad de Granger* a cada una de las variables explicativas, siendo las ventas minoristas la única variable que rechaza la hipótesis nula, y que por lo tanto puede considerarse como endógena.

De todas maneras, existen fuertes argumento teóricos para afirmar que esta última no es la única variable endógena, ya que, lo más probable, es que exista una retroalimentación principalmente entre las variables IMACEC, Ventas INE, Ventas Minoristas y Producción del INE, por lo que se debiera utilizar alguna metodología que permitieran sortear este inconveniente.

Para lidiar con este problema se debe incorporar explícitamente en la estructura la relación entre las variables. Con este fin se estiman dos modelos VAR alternativos en los cuales se posibilita la interacción entre las variables.

En el primer modelo VAR se incluyen como variables endógenas C12IMA, C12OCUPA (variación en doce meses del nivel de ocupación) y C12VTAMINOR (Variación en doce meses en las ventas minoristas). Además se incluyen como variables exógenas CDIAL, C12TERM, C12MINAS, C12_IMP, todas anteriormente

definidas y además se agrega la variable C12_CEME, el cambio a doce meses en los despachos de cemento. Es importante aclarar que las importaciones no son necesariamente exógenas al sistema, pero se ha optado por esta alternativa para aprovechar la prontitud con la que este indicador se conoce.

Como se observa en la tabla N° 6, el VAR estimado incluye tres rezagos, los cuales se han elegido utilizando dos criterios. Primero, el criterio de minimización de error de predicción final, y segundo, que al mismo tiempo los residuos resultantes para cada ecuación se aproximen lo más posible a un ruido blanco, es decir, que no muestren correlación con los valores pasados y que la distribución de los mismos sea normal.

Tabla N° 6. VAR 1(*)

	C12IMA	C12OCUPA	C12VTAMINOR		C12IMA	C12OCUPA	C12VTAMINOR
C12IMA(-1)	0.283162 (0.05743) -493.087	0.078452 (0.03458) -226.843	0.971224 (0.26828) -362.019	C12VTAMINOR(3)	-0.022704 (0.01987) (-1.14256)	-0.013151 (0.01197) (-1.09892)	-0.033438 (0.09283) (-0.36020)
C12IMA(-2)	0.163722 (0.06117) -267.653	0.032231 (0.03684) (0.87493)	-0.312807 (0.28877) (-1.09462)	Constante	0.008425 (0.00233) -361.297	-0.000893 (0.00140) (-0.63601)	-0.005033 (0.01089) (-0.46206)
C12IMA(-3)	0.049190 (0.06169) (0.79747)	-0.019671 (0.03715) (-0.52950)	-0.304770 (0.28819) (-1.05755)	C12DIAL	0.001985 (0.02287) (0.08699)	0.002161 (0.01374) (0.15728)	-0.190781 (0.10658) (-1.78999)
C12OCUPA(-1)	0.083534 (0.15483) (0.53952)	1.029521 (0.09324) -110.411	-0.508829 (0.72333) (-0.70304)	C12TERM	0.019820 (0.00407) -494.322	0.000250 (0.00247) (0.10357)	0.010805 (0.01873) (0.57684)
C12OCUPA(-2)	-0.366397 (0.22973) (-1.59490)	-0.025496 (0.13835) (-0.18428)	0.003358 -107.324 (0.00313)	C12MINAS	0.057893 (0.01425) -476.514	-0.007967 (0.00888) (-0.92857)	0.100784 (0.06656) -151.414
C12OCUPA(-3)	0.308824 (0.15828) -197.605	-0.195305 (0.09412) (-2.07507)	0.274571 (0.73017) (0.37607)	C12PROD_INE	0.209787 (0.04117) -609.573	0.013441 (0.02479) (0.54211)	0.010897 (0.19233) (0.05666)
C12VTAMINOR(-1)	0.008453 (0.02049) (0.41303)	0.002484 (0.01234) (0.20128)	0.487083 (0.09573) -608.830	C12VTAS_INE	0.110012 (0.03107) -354.728	0.041359 (0.01868) -221.443	0.031959 (0.14488) (0.22059)
C12VTAMINOR(-2)	-0.008979 (0.02287) (-0.39887)	-0.006162 (0.01356) (-0.45452)	0.251472 (0.10516) -239.127	C12GPA	-0.012682 (0.00545) (-232573)	-0.005249 (0.00328) (-1.59847)	0.004499 (0.02547) (0.17662)
				C12_IMP	0.052231 (0.00933) -666.977	-0.012753 (0.00962) (-2.26953)	-0.051918 (0.04359) (-1.42053)
				R2	0.941671	0.933550	0.653134
				R2-Ajustado	0.881742	0.922286	0.594054

(*) El Conjunto de datos engloba el período que va desde enero de 1992 hasta julio de 2002.

En el segundo VAR se incluyen como variables endógenas C12IMA, C12OCUPA, C12_IMP y C12PROD_INE (Variación en doce meses del índice de producción del INE). Además se incluyen como variables exógenas CDIAL, C12TERM,

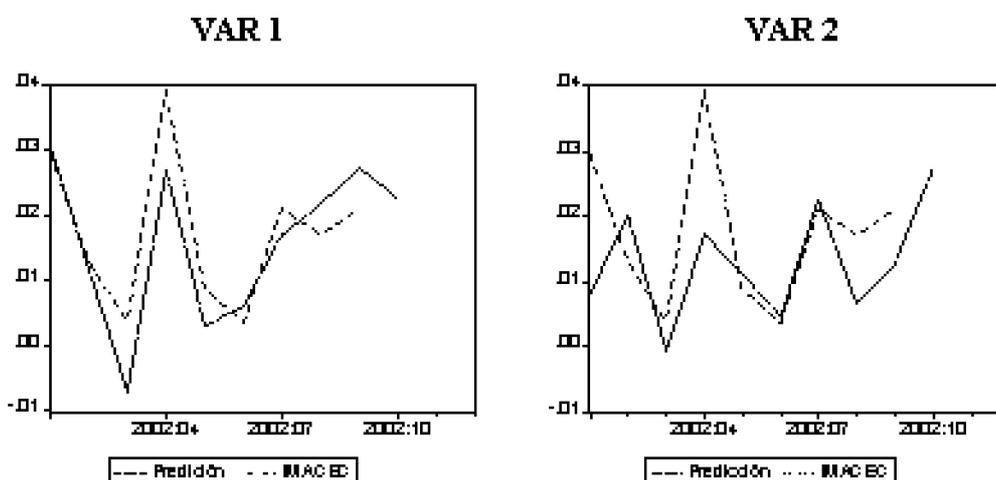
C12MINAS, C12_IMP, C12VTAMINOR, C12_CEME, TCO_DFLAC, todas anteriormente definidas y además se agrega la variable C12TPM, el cambio a doce meses en la tasa de política monetaria. El modelo estimado se presenta a continuación.

Tabla N° 7. VAR 2

	C12IMA	C12OCUPA	C12_IMP	C12PROD_INE
C12IMA(1)	0.310803 (0.15212) [2.04312]	0.045924 (0.05363) [0.72173]	-0.057936 (1.02428) [0.05698]	-0.281418 (0.27990) [1.00542]
C12IMA(2)	0.226898 (0.13637) [1.65432]	-0.059589 (0.05706) [1.04468]	0.181009 (0.91821) [0.19713]	0.167172 (0.25092) [0.66629]
C12OCUPA(1)	0.380531 (0.23751) [1.60259]	0.923092 (0.09935) [9.29170]	2.943184 (1.59922) [1.84039]	0.390765 (0.43701) [0.89417]
C12OCUPA(2)	-0.288546 (0.21408) [1.34833]	-0.191460 (0.08954) [2.13815]	-2.938787 (1.44145) [2.03877]	-0.205630 (0.39390) [0.52204]
C12_IMP(1)	0.006997 (0.01735) [0.34537]	-0.007369 (0.00725) [1.01460]	0.275529 (0.11692) [2.35651]	0.009651 (0.03195) [0.30205]
C12_IMP(2)	0.023891 (0.01710) [1.39724]	-0.001563 (0.00715) [0.21895]	0.110221 (0.11513) [0.95737]	0.019940 (0.03145) [0.63379]
C12PROD_INE(1)	-0.025888 (0.05453) [0.40120]	0.025965 (0.02699) [0.96167]	0.022988 (0.43447) [0.05291]	0.282280 (0.11873) [2.37759]
C12PROD_INE(2)	-0.057186 (0.05224) [0.91880]	0.072651 (0.02603) [2.79065]	0.094854 (0.41908) [0.22636]	-0.068700 (0.11452) [0.99989]
Constante	0.016679 (3.06677)	-0.000842 (0.36999)	0.091962 (2.51045)	0.019011 (1.89920)
C12DIAL	0.104550 (0.03053) [3.42475]	0.009678 (0.01277) [0.75791]	0.296030 (0.20555) [1.44017]	0.312121 (0.05617) [5.56669]
C12TERM	0.012341 (0.00673) [1.83362]	0.004618 (0.00282) [1.64031]	-0.046685 (0.04532) [1.03013]	-0.007272 (0.01235) [0.58717]
C12MINAS	0.052454 (0.02649) [1.98079]	-0.003975 (0.01105) [0.35879]	0.053603 (0.17834) [0.30057]	-0.029326 (0.04873) [0.60175]
C12VTAMINOR	0.013147 (0.03897) [0.33740]	0.028172 (0.01630) [1.72846]	-0.347498 (0.25237) [1.32447]	0.040638 (0.07170) [0.56680]
C12TPM	-0.018513 (0.00723) [2.57434]	0.000549 (0.00302) [0.18159]	-0.071672 (0.04868) [1.47221]	-0.034334 (0.01330) [2.58081]
C12_CEME	0.082574 (0.01716) [4.81229]	0.021014 (0.00715) [2.92783]	0.461015 (0.11554) [3.99021]	0.130042 (0.03157) [4.11888]
TCO_DFLAC	-0.096559 (0.03223) [2.99636]	0.013858 (0.01345) [1.02806]	-1.121353 (0.21698) [5.16791]	-0.137589 (0.05929) [2.32045]

C-12IG PA	0.013711 (0.00824) [1.66458]	-0.006852 (0.00346) [-1.96891]	0.051964 (0.05546) [0.92976]	0.040401 (0.01516) [2.66587]
R-squared	0.890341	0.897111	0.803533	0.714578
Adj. R-squared	0.869700	0.877744	0.766551	0.660851

Gráfico N° 2 Predicciones Año 2002



III. Elección del Predictor

En la Tabla N° 8 se presentan los valores predichos por los diferentes modelos y los observados hasta Julio de 2002. También se adjunta un promedio de los valores predichos, y en las dos últimas filas el error promedio y la varianza del error. El modelo óptimo sería aquel que muestre el menor error promedio en valor absoluto, en lo posible cero, y al mismo tiempo la menor varianza. Curiosamente el modelo con menor error medio es el modelo de confección más sencilla, el ARMA, con un error de 0,07%, en tanto que el mayor error promedio lo exhibe el primer modelo VAR con 0,48%, aunque en este caso con sesgo negativo.

Tabla N° 8

Mes	Modelo			Modelo con Eógenas	Modelo VAR 1	Modelo VAR 2	Promedio Predicción
	VERDADERO	ARMA(1,3)	Líder				
Enero	0,0296	0,0144	0,0216	0,0373	0,0398	0,0211	0,0269
Febrero	0,0131	0,0272	0,0238	0,0135	0,0105	0,0152	0,0180
Marzo	0,0039	0,0219	0,0069	-0,0060	-0,0043	0,0020	0,0041
Abril	0,0391	0,0150	0,0395	0,0343	0,0216	0,0282	0,0277
Mayo	0,0088	0,0232	0,0235	0,0083	0,0015	0,0085	0,0130
Junio	0,0035	0,0007	0,0094	0,0248	0,0002	0,0039	0,0078
Julio	0,0210	0,0192	0,0440	0,0285	0,0173	0,0385	0,0262
Error medio		0,0007	0,0045	0,0024	-0,0048	-0,0032	-0,0001
Desviación Estándar		0,0176	0,0080	0,0110	0,0091	0,0052	0,0053

Sin embargo, si se observa la varianza del error se destaca que el modelo ARMA también muestra la mayor varianza del error de predicción, casi un 1,8%. Por ejemplo, dado que el valor predicho para Julio fue de 1,9% con ese modelo, un intervalo de confianza del 90% incluiría los valores que van desde -1,7 hasta 5,5%, lo que claramente representa un intervalo demasiado amplio. Bajo este criterio, el modelo elegido, y por lo tanto con menor varianza, es el VAR 2, con una desviación estándar de 0,5%. Sin embargo, como se desprende de la última columna, el indicador que posee las mejores propiedades estadísticas es el promedio de las proyecciones, con un error promedio de sólo 0,01% y una desviación estándar de 0,6%, la segunda menor después del VAR 2.

IV. Resumen y Conclusiones

En el presente trabajo se han explorado una amplia gama de modelos propuesto por la teoría para la proyección de variables, desde modelos uniecuacionales, hasta modelos más ricos de ecuaciones simultaneas. Todos ellos predicen de manera razonable los valores observados para el presente año. Sin embargo, de entre ellos se destacan el VAR 2 y el promedio de las predicciones arrojadas por los diferentes modelos. No obstante, la proyección más confiable sería la obtenida por el promedio de las proyecciones, ya que, además de sus buenas propiedades estadísticas, ya que corrige por posibles sesgos en la estimación mediante los VAR, es la proyección que incorpora mayor cantidad de información. Esta contiene información respecto a la historia de las variables, respecto a señales que puedan dar los valores pasados de otras variables relacionadas, además de incluir de alguna manera la información nueva a través del indicador coincidente.

No obstante, al momento de realizar estimaciones para plazos más prolongados se debe confiar solamente en los modelos que no incluyen variables exógenas contemporáneas, es decir, el ARMA, el Líder y el VAR sin exógenas.

Por último, la importante tarea de proyección de las variables económicas más relevantes es un proceso continuo de prueba y error, todos los modelos son perfeccionables a medida que aparece nueva información. Es por esto que esperamos que este sea sólo un punto de partida que se irá perfeccionando de manera permanente.

Bibliografía

Bravo, Felipe. (2002). "El Pass Throw del Tipo de Cambio en Chile". Documentos de Trabajo N°. Banco Central de Chile. Julio de 2002.

Bravo, Felipe; Franken, Helmut. (2001). "Un Indicador Líder del IMACEC". Documentos de Trabajo N° 99. Banco Central de Chile. Junio 2001.

Enders, Walter. (1995). "Applied Econometric Time Series". New York. John Wiley and Sons.

Green, William. (1998). "Análisis Econométrico". Tercera Edición. Prentice Hall.

Hamilton, James. (1994). "Time Series Analysis". Princeton University Press.

Harvey, Adrew. (1989). "Forecasting, Structural Time Series Models and de Kalman Filter". Cambridge University Press.