

Banco Central de Chile
Documentos de Trabajo

Central Bank of Chile
Working Papers

N° 620

Marzo 2011

**UN TEST CONJUNTO DE SUPERIORIDAD
PREDICTIVA PARA LOS PRONÓSTICOS DE
INFLACIÓN CHILENA**

Pablo Pincheira B.

La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: <http://www.bcentral.cl/esp/estpub/estudios/dtbc>. Existe la posibilidad de solicitar una copia impresa con un costo de \$500 si es dentro de Chile y US\$12 si es para fuera de Chile. Las solicitudes se pueden hacer por fax: (56-2) 6702231 o a través de correo electrónico: bcch@bcentral.cl.

Working Papers in PDF format can be downloaded free of charge from: <http://www.bcentral.cl/eng/stdpub/studies/workingpaper>. Printed versions can be ordered individually for US\$12 per copy (for orders inside Chile the charge is Ch\$500.) Orders can be placed by fax: (56-2) 6702231 or e-mail: bcch@bcentral.cl.



BANCO CENTRAL DE CHILE

CENTRAL BANK OF CHILE

La serie Documentos de Trabajo es una publicación del Banco Central de Chile que divulga los trabajos de investigación económica realizados por profesionales de esta institución o encargados por ella a terceros. El objetivo de la serie es aportar al debate temas relevantes y presentar nuevos enfoques en el análisis de los mismos. La difusión de los Documentos de Trabajo sólo intenta facilitar el intercambio de ideas y dar a conocer investigaciones, con carácter preliminar, para su discusión y comentarios.

La publicación de los Documentos de Trabajo no está sujeta a la aprobación previa de los miembros del Consejo del Banco Central de Chile. Tanto el contenido de los Documentos de Trabajo como también los análisis y conclusiones que de ellos se deriven, son de exclusiva responsabilidad de su o sus autores y no reflejan necesariamente la opinión del Banco Central de Chile o de sus Consejeros.

The Working Papers series of the Central Bank of Chile disseminates economic research conducted by Central Bank staff or third parties under the sponsorship of the Bank. The purpose of the series is to contribute to the discussion of relevant issues and develop new analytical or empirical approaches in their analyses. The only aim of the Working Papers is to disseminate preliminary research for its discussion and comments.

Publication of Working Papers is not subject to previous approval by the members of the Board of the Central Bank. The views and conclusions presented in the papers are exclusively those of the author(s) and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.

Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180, Santiago, Chile
Teléfono: (56-2) 3882475; Fax: (56-2) 3882231

UN TEST CONJUNTO DE SUPERIORIDAD PREDICTIVA PARA LOS PRONÓSTICOS DE INFLACIÓN CHILENA

Pablo Pincheira B.
Banco Central de Chile

Resumen

Diversos pronósticos de inflación son habitualmente construidos en distintos momentos del tiempo y en base a conjuntos de información diferentes. Sería esperable que un pronóstico construido sobre la base de un cierto conjunto de información fuese más preciso que otro construido en base a un conjunto de información anidado en el primero. Para evaluar esta hipótesis, el presente trabajo aplica un test conjunto de superioridad predictiva a las proyecciones de inflación Chilena disponibles a través de las siguientes fuentes: Bloomberg, Consensus Economics, la Encuesta de Expectativas Económicas del Banco Central de Chile y un promedio de modelos ESARIMA. Los resultados indican que en términos de Error Cuadrático Medio no es posible rechazar la hipótesis nula de optimalidad para el perfil conjunto de proyecciones. No obstante, cuando el test es aplicado a las Proyecciones al Cuadrado, en línea con la propuesta de Patton y Timmermann (2010), el test conjunto rechaza la hipótesis de optimalidad. El análisis de estos resultados sugiere que existe una violación de una condición de ortogonalidad implicada por la optimalidad bajo pérdida cuadrática, y que esta violación es transversal al conjunto de proyecciones bajo consideración y al horizonte predictivo, todo lo cual indica la existencia de espacio para perfeccionar el perfil de proyecciones.

Abstract

Optimality under quadratic loss implies that forecasts built using a large information set should perform at least as well as forecasts built using a more restricted and nested information set. In this article we use a joint test of superior predictive ability to test this optimality condition for the term structure of several Chilean inflation forecasts coming from the following sources: Bloomberg, Consensus Economics, the Survey of Professional Forecasters and an average of selected seasonal univariate models. We do this by taking advantage of the fact that these sets of forecasts are built at different moments in time and, more importantly, using different and nested information sets. Our results indicate that the null hypothesis of optimality under quadratic loss cannot be rejected when Mean Squared Error is used to evaluate the term structure of the forecasts. Nevertheless, when the joint test is carried out to evaluate the term structure of the Mean Squared Forecasts, as suggested by Patton and Timmermann (2010), the joint test rejects the null hypothesis of optimality. Further analysis of our results reveals that this rejection is associated with a violation of an orthogonality condition that should be satisfied when forecasts are optimal. Moreover, this violation seems to stand both across different sources of forecasts and across different forecasting horizons. This suggests that there is room for improvement in the term structure of Chilean inflation forecasts.

la sociedad. En el trabajo se demuestra que, si este costo es cero, el impuesto inflacionario óptimo es cero. Sin embargo, si el costo de producir depósitos y el gasto del gobierno son positivos, la tasa de inflación óptima es positiva. Inclusive, la tasa de inflación óptima es una función creciente del gasto del gobierno y el costo de producir depósitos.

* * Proyecto FONDECYT 11080109. Agradezco especialmente a Carlos Medel por su brillante colaboración y asistencia. Las ideas y opiniones presentadas en este trabajo no necesariamente representan la visión del Banco Central de Chile o sus autoridades. Los errores u omisiones son responsabilidad exclusiva del autor.

** Dirección: Agustinas 1180, Santiago, Chile. Teléfono: (562) 6702874. E-mail: ppinchei@bcentral.cl.cl.

1. Introducción

Cuando coexisten diversos pronósticos para una misma variable económica, resulta natural investigar la precisión relativa entre ellos. Esto tradicionalmente es realizado a través de múltiples comparaciones bilaterales del Error Cuadrático Medio (ECM) de proyección. Al momento de hacer inferencia, la manera usual de proceder consiste en formular una serie de hipótesis de igualdad predictiva entre dos pares de métodos o modelos predictivos. Luego es aplicado algún test como el de Diebold y Mariano (1995) y West (1996) o el de Giacomini y White (2006). Ejemplos en los que estas estrategias han sido utilizadas son los artículos de Pincheira y García (2009), Andersson et al. (2007), Groen et al. (2009), Ang et al. (2009), Pincheira y Alvarez (2009) y Pincheira (2010), por solo citar algunos trabajos.

Esta tradicional metodología para comparar la capacidad predictiva entre varios modelos presenta la natural complejidad asociada a la formulación de hipótesis múltiples, que se refiere a la dificultad de mantener bajo control el error de tipo I del test¹. Sin la formulación de un contraste conjunto que específicamente esté diseñado para tener un error de tipo I no superior a un nivel α , la aplicación de repetidas pruebas bilaterales no necesariamente garantizará el nivel del error de tipo I deseado. Una manera de paliar este problema hace uso de las llamadas cotas de Bonferroni. En términos sencillos esta metodología plantea transformar una dócima de hipótesis conjunta de tamaño α , en N pruebas de tamaño α/N que solo involucran a hipótesis univariadas. No obstante la facilidad de implementación de esta estrategia, en general es criticada por poder llegar a ser en extremo conservadora y sin una capacidad real de rechazar la hipótesis nula. Esto es mencionado en el trabajo de Patton y Timmermann (2010) y comprobado con simulaciones en las secciones finales de este documento.

En forma adicional a la dificultad planteada en el párrafo anterior, la evaluación predictiva de múltiples series de pronósticos puede ser más compleja aún cuando las proyecciones son realizadas en distintos momentos del tiempo. En ese escenario es esperable que las proyecciones elaboradas con mayor información tengan un mejor desempeño que aquellas construidas con menor información. Paralelamente, también debería ser esperable que, bajo ciertos supuestos, los pronósticos a horizontes cortos tuviesen un desempeño mejor que aquellos diseñados para horizontes más largos. De hecho, Patton y Timmermann (2010) muestran que proyecciones óptimas bajo pérdida cuadrática deberían seguir un comportamiento de esta naturaleza. En este sentido, surge natural la posibilidad de evaluar la optimalidad de un conjunto de proyecciones, analizando la distancia que existe entre el comportamiento efectivo de ellas y el comportamiento esperado de un pronóstico óptimo. Cuando la teoría econométrica sugiere que un cierto pronóstico debiese ser mejor que otro, cobra sentido plantear una hipótesis nula de superioridad de habilidad predictiva. Algunos tests de esta naturaleza han sido presentados por White (2000), Hansen (2005) y recientemente también por Patton y Timmerman (2010), quienes utilizan implícitamente tests de superioridad predictiva basados en los trabajos de Wolak (1987,1989).

Interesantemente, para el caso chileno existen diversas proyecciones de inflación publicadas periódicamente en distintos momentos del tiempo por distintas entidades y analistas. Dichas proyecciones corresponden a las de Bloomberg, Consensus Economics, y la Encuesta de Expectativas Económicas (EEE) realizada por el Banco Central de Chile (BCCh). Adicionalmente siempre es posible contar con proyecciones de simples modelos de series de tiempo univariados como las presentadas en Pincheira y García (2009). Al requerir solo de los rezagos de la inflación, estos simples modelos utilizan un conjunto de información muy pequeño, que supondremos está contenido en los conjuntos utilizados por los analistas que contemplan las diversas encuestas mencionadas anteriormente.

Frente a este perfil de proyecciones de inflación construidas en base a conjuntos de información que supondremos anidados, cabe naturalmente hacerse la pregunta acerca de la superioridad predictiva de aquellas proyecciones construidas sobre la base de mayor información. Paralelamente, y en el caso de proyecciones

¹El error de tipo I corresponde a la probabilidad de rechazar la hipótesis nula cuando ella es verdadera.

a varios distintos horizontes, también surge de manera natural evaluar la superioridad de las proyecciones construidas para horizontes más cercanos. Es este precisamente el principal objetivo de este trabajo: evaluar si el perfil de proyecciones de inflación disponible para Chile es coherente con las propiedades de optimalidad bajo pérdida cuadrática, en el sentido de entregar predicciones de mejor calidad en la medida que se cuenta con más información disponible, o en la medida que el horizonte predictivo sea menor. Para llevar a cabo esta tarea, y evitar los problemas asociados al correcto control del error de tipo I en pruebas múltiples, construimos dos tests basados en los estadísticos del máximo y mínimo de un conjunto de variables aleatorias. Tests de esta naturaleza han sido desarrollados independientemente por White (2000) y Pincheira (2006) en contextos algo distintos al que nos ocupa en el presente documento. Hansen (2005) presenta explícitamente un test de superioridad de capacidad predictiva basándose en una estudentización del test propuesto por White (2000) y de alguna manera, reinterpreta el test de White como uno de superioridad de capacidad predictiva.

La construcción de los tests se hace con el propósito de contrastar cuatro propiedades. Primero se evalúa si el perfil de proyecciones de inflación para un horizonte fijo presenta una precisión no decreciente con el tamaño del conjunto de información utilizado para proyectar. Esto significa que el ECM de las proyecciones construidas en el instante t_0 para el horizonte T no debiese ser menor que el ECM de las proyecciones construidas en t_1 cuando $t_0 < t_1$, y el horizonte predictivo sigue siendo T . Un segundo test es construido para evaluar que el Pronóstico Cuadrático Medio (PCM) de las proyecciones también presente un patrón no decreciente con el tamaño del conjunto de información utilizado para proyectar a un horizonte fijo T . Esta propiedad ha sido introducida recientemente por Patton y Timmermann (2010) como una característica adicional que debe presentar un perfil de proyecciones óptimo bajo pérdida cuadrática. En tercer lugar consideramos proyecciones a distintos horizontes pero construidas en el mismo instante t . En este contexto evaluamos que el ECM de las proyecciones a horizontes más distantes sea mayor o igual que el ECM de proyecciones a horizontes cortos. Finalmente, también evaluamos si el PCM de las proyecciones a horizontes largos es menor o igual que el PCM de las proyecciones de corto plazo.

Nuestros resultados indican que en términos de ECM no es posible rechazar la hipótesis nula de optimalidad del perfil de proyecciones considerado. Esto ocurre tanto cuando evaluamos las proyecciones a horizonte fijo como cuando las evaluamos a horizonte variable. Similarmente, la hipótesis nula de optimalidad tampoco puede ser rechazada cuando evaluamos los PCM a horizonte variable. Sin embargo, el test conjunto para los PCM rechaza la hipótesis de optimalidad cuando se consideran las proyecciones a horizonte fijo, sugiriendo la existencia de espacio para perfeccionar el perfil de las proyecciones.

Más allá de los resultados empíricos que obtenemos en el artículo, creemos que estamos haciendo una contribución en dos dimensiones. En primer lugar, y dentro de los límites de la literatura revisada, pensamos que este es el primer artículo en testear conjuntamente el perfil no creciente de los ECM para un conjunto de proyecciones realizadas en distintos momentos del tiempo. En segundo lugar, no conocemos de ningún otro artículo que evalúe para Chile el ordenamiento esperado que los PCM de las proyecciones de inflación debiesen presentar si los pronósticos fueran óptimos. Para terminar, es importante señalar que los tests aquí utilizados pueden ser directamente aplicados a evaluar otras hipótesis conjuntas que involucren la comparación de múltiples series de pronósticos. Esto es muy relevante, toda vez que es habitual en la literatura comparar la capacidad predictiva de una determinada serie de pronósticos, con respecto a la pléyade de otras proyecciones disponibles en el mercado o en la literatura.

El resto del trabajo se organiza de la siguiente manera. En la siguiente sección se revisa brevemente la literatura relacionada. Luego, en la sección 3 se entregan algunas características de los datos utilizados, para después en la sección 4 resumir la derivación de las propiedades teóricas que son contrastadas en la sección 5. En la sección 6 se hace un análisis de los resultados empíricos para finalmente concluir en la sección 7.

2. Revisión de la Literatura

Existen numerosos trabajos que evalúan las proyecciones públicas de diversas variables económicas como crecimiento del PIB e inflación entre otras. Entre ellos destacamos las recientes contribuciones de Croushore (2010), Capistrán y López-Moctezuma (2010), Groen et al (2009), Pincheira (2010), así como el ya más tradicional artículo de Ang, Bekaert y Wei (2007). En general, en casi ninguno de estos artículos se considera la dinámica temporal en la construcción de pronósticos, siendo la única excepción por nosotros conocida la del artículo de Pincheira y Alvarez (2009). A pesar de plantear explícitamente una hipótesis conjunta de superioridad de capacidad predictiva, dicho artículo no plantea un test conjunto para enfrentar la dócima de hipótesis y desde ese punto de vista no es claro que se realice un control adecuado del error de tipo I. Así, la práctica mayoritaria plantea hipótesis de igualdad de capacidad predictiva por sobre la hipótesis de superioridad predictiva considerada en este trabajo. Además el enfoque generalmente utilizado se refiere a la comparación de dos modelos rivales, a pesar de disponer de más de dos fuentes para evaluación.

Otros trabajos evalúan pronósticos a través de sus propiedades de optimalidad bajo alguna función de pérdida, como las descritas en Elliot y Timmermann (2008). Llama fuertemente la atención que una gran mayoría de esta literatura encuentre que diversas series de proyecciones no satisfacen propiedades provenientes de pronósticos óptimos bajo pérdida cuadrática. Entre estos estudios destaca el trabajo de Joutz y Stekler (2000), quienes analizando las predicciones de la Reserva Federal de Estados Unidos, encuentran que la serie de errores de los pronósticos de inflación muestra excesos de autocorrelación. Más recientemente, Capistrán y López-Moctezuma (2010) encuentran que los pronósticos de consenso de los especialistas en México, en general, no superan pruebas de ausencia de sesgo y de ausencia de autocorrelación en exceso, lo que sugiere oportunidades para mejorar dichos pronósticos. Para el caso de la Zona Euro, Bowls, Friz, Genre, Kenny, Meyler y Rautanen (2007) analizan las predicciones económicas realizadas por una serie de analistas privados. Ellos encuentran, entre otras cosas, que para el periodo que va desde el primer trimestre de 1999 hasta el último trimestre del 2006, los analistas han tendido a subestimar significativamente la inflación y a sobrestimar el crecimiento económico. Otro interesante artículo que analiza la calidad de predicciones económicas es el de Loungani (2001) quien considera las proyecciones de crecimiento económico para una serie de países industrializados y en desarrollo durante el periodo 1989 - 1998. Las proyecciones son obtenidas de Consensus Economics y los resultados muestran señales de suboptimalidad tanto en los pronósticos de países industrializados como en los de países en desarrollo.

Para el caso de Chile, se destacan los trabajos de Chumacero (2001), el cual documenta una subestimación sistemática de las proyecciones privadas del PIB durante 1986–1998, el de Bentancor y Pincheira (2010) donde se analizan los resultados para la inflación de la EEE, encontrando un significativo sesgo a la subestimación y exceso de autocorrelación en la segunda mitad del período muestral evaluado. Al corregir el sesgo y exceso de autocorrelación, en un ejercicio fuera de muestra, los autores reducen el ECM y el sesgo de manera significativa. También en esta línea de análisis se encuentra el trabajo de Pincheira y Alvarez (2009). Los autores evalúan las proyecciones de inflación provenientes de Bloomberg, Consensus Economics, la EEE del BCCh, proyecciones auxiliares del BCCh y modelos de series de tiempo. Uno de los objetivos de ese trabajo, que coincide con el del presente trabajo, es evaluar la superioridad predictiva de los pronósticos construidos sobre la base de mayor información. Es importante recalcar que en Pincheira y Alvarez no se realiza un test conjunto propiamente tal sino que una secuencia de tests bilaterales de igualdad de capacidad predictiva, con lo que no es claro para el lector el adecuado control del error de tipo I que se haya tenido en ese artículo. Con todo, ellos concluyen que no es posible rechazar la hipótesis nula de menor o igual ECM mientras mayor sea el conjunto de información utilizado en la proyección. En términos de otras consideraciones, los autores advierten un cierto nivel de suboptimalidad en algunas de las proyecciones analizadas, especialmente en los horizontes de predicción más largos, lo que sugiere la existencia de espacio para mejorar la precisión de varios pronósticos.

La primera parte del presente trabajo sigue de cerca lo presentado en Pincheira y Alvarez (2009), aunque se destaca que la principal diferencia respecto al trabajo de esos autores es que en este artículo se realiza la

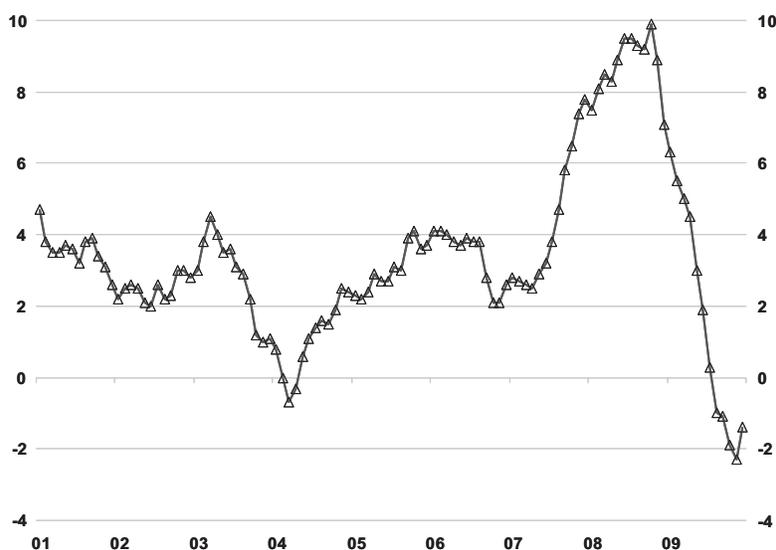
prueba para el perfil conjunto, de acuerdo al planteamiento de White (2000) siguiendo algunas modificaciones sugeridas por Hansen (2005) y Romano y Wolf (2005), para tener un control adecuado del error de tipo I. Además el presente trabajo no solo se centra en el perfil conjunto de ECM sino que añade el perfil de PCM en línea con el artículo de Patton y Timmermann (2010).

De acuerdo a esta breve y selecta revisión de la literatura, se infiere que un alto porcentaje de las proyecciones de variables macro están sujetas a ineficiencias, y que pocos trabajos enfocan la evaluación de proyecciones de manera conjunta y considerando la dinámica temporal en la elaboración de las proyecciones, un aspecto distintivo del presente trabajo. En las siguientes secciones se observará cómo nuestros resultados se ajustan a los hallazgos de la literatura relacionada.

3. Datos

El análisis se realiza con datos chilenos de frecuencia mensual. La fuente de los datos efectivos es el Instituto Nacional de Estadísticas (INE), mientras que las proyecciones provienen de Bloomberg, Consensus Economics, la EEE del BCCh, y una combinación de modelos de series tiempo presentados en Pincheira y García (2009). Todas las series tienen frecuencia mensual y están denominadas en porcentaje, y representan la variación anual del IPC. Para algunas pruebas y estadísticos descriptivos se considera la mayor muestra disponible sin observaciones faltantes de cada proyección, mientras que el test conjunto y otros análisis se realizan con una muestra balanceada desde Diciembre del 2003 hasta Diciembre del 2009 (73 observaciones). En el Gráfico 1 se presenta la serie original de inflación desde Enero del 2001 hasta Diciembre del 2009. Se advierte un importante crecimiento de la inflación desde mediados del 2007 hasta fines del 2009.

Gráfico 1: Inflación del IPC en Chile: 2001–2009 (*).



(*) Variación Porcentual en 12 Meses. Fuente: INE.

3.1. Bloomberg

Las proyecciones que denominamos bajo el nombre “Bloomberg” corresponden a la media de las respuestas de la encuesta efectuada por Bloomberg a distintos analistas e instituciones financieras. Se dispone de una muestra que abarca desde Enero del 2001 hasta Diciembre del año 2009. Estas proyecciones se dan a conocer aproximadamente una semana antes de conocer el dato efectivo de inflación y de hecho, dentro de todas las proyecciones que consideramos, son las últimas en darse a conocer antes de la publicación del dato efectivo. El horizonte de proyección de esta encuesta es de un mes, aunque en términos efectivos tiene un horizonte de alrededor de una semana. Existe una observación faltante (Noviembre 2003) la que no recibe

un tratamiento especial, y es la que condiciona la selección de la muestra alternativa que se utiliza para los test de superioridad predictiva.

3.2. Consensus Economics

Las proyecciones que denominamos bajo el nombre “Consensus” corresponden a la media de las respuestas de la encuesta efectuada por Consensus Economics a distintos analistas. La encuesta de Consensus se realiza alrededor de tres semanas antes de conocer los datos efectivos de inflación. Esta encuesta provee proyecciones en horizontes de uno a seis meses. La muestra existente considera datos desde Julio de 1993 hasta la actualidad. Sin embargo, se disponen de datos sin observaciones faltantes solo desde Abril del 2001 en adelante para los seis horizontes de proyección.

3.3. Encuesta de Expectativas Económicas

La EEE es levantada por el BCCh a comienzos de cada mes, y entrega proyecciones de inflación a uno, tres y doce meses hacia adelante.² Se aplica a un selecto grupo de académicos, consultores, ejecutivos y asesores del sector privado, quienes informan, además de sus proyecciones de inflación, pronósticos de otras variables económicas de interés. Dado que las proyecciones individuales de cada analista no son publicadas por el BCCh, utilizamos la mediana de las proyecciones. Este dato es de público acceso a través de la página web del BCCh. La EEE comenzó a operar desde el año 2000 y en varias oportunidades se modificó su contenido. Salvo cambios menores, ha permanecido con pocas alteraciones desde Noviembre del 2004. Se cuenta con una muestra sin observaciones faltantes desde Enero del 2001 hasta la actualidad para pronósticos un mes hacia adelante, desde Noviembre del 2001 para pronósticos tres meses hacia adelante y desde Septiembre del 2001 para doce meses hacia adelante. En promedio, durante el periodo 2000–2009, 35 personas contestaron el cuestionario cada mes³.

3.4. Proyección ESARIMA

Corresponde al promedio aritmético de los pronósticos provenientes de ocho modelos de series de tiempo de naturaleza similar. Dichos modelos se describen en Pincheira y García (2009) y corresponden a distintas especificaciones de modelos ESARIMA (*Extended Seasonal ARIMA*) que se pueden anidar en la siguiente expresión:

$$\pi_t - \pi_{t-1} = \delta + \rho(\pi_{t-1} - \pi_{t-2}) + \varepsilon_t - \theta\varepsilon_{t-1} + \theta_E\varepsilon_{t-12} + \theta_3\varepsilon_{t-13},$$

donde π_t es la inflación en variación a 12 meses en t , ε_t es un término de error, y δ , ρ , θ , θ_E y θ_3 son parámetros a estimar (E es la frecuencia de los datos, que para el caso mensual corresponde a $E = 12$). Los ocho modelos se generan por la combinatoria de casos en que δ , ρ , y θ_3 son estimados o restringidos a cero. Notemos que cuando θ_3 es estimado, su estimación es restringida a satisfacer la condición $\hat{\theta}_3 = \hat{\theta}\hat{\theta}_E$. Estos ocho modelos siempre incluyen los términos que acompañan a los parámetros θ y θ_E . Con estos modelos se construyen pronósticos para siete distintos horizontes: de uno a seis meses y un año hacia adelante. La muestra para los siete horizontes considerados comienza en Enero del 2001 y finaliza en Diciembre del 2009.

La Tabla 1 resume los períodos muestrales que cubre cada fuente de proyecciones considerada. En la Tabla 2 se presentan algunos estadísticos descriptivos de las series tanto para la muestra completa como para la submuestra utilizada en los tests de superioridad predictiva (Diciembre del 2003 – Diciembre del 2009). Se observa que, en general, las proyecciones presentan una estimación del error cuadrático medio relativamente bajo a horizontes menores a un año. Recordemos que el $ECM(h)$ se define por:

$$ECM(h) = \mathbb{E}(Y_{t+h} - y_t^f(h))^2$$

²En estricto rigor, la EEE no pregunta explícitamente por proyecciones tres meses hacia adelante. Simplemente pregunta por proyecciones un mes hacia adelante y por la inflación promedio de los siguientes dos meses. La proyección a tres meses que usamos en este artículo proviene de aplicar la aproximación logarítmica que indica que la inflación acumulada en tres meses es la suma de las inflaciones mensuales.

³Una descripción técnica de la EEE y evaluación de sus resultados se encuentra en Pedersen (2010).

en que $y_t^f(h)$ es la proyección de Y_{t+h} para h períodos adelante. La raíz del ECM se denota por $RECM$ y corresponde a la raíz cuadrada del ECM. Al ser un momento poblacional, el ECM debe ser estimado. Un estimador natural está dado por:

$$\widehat{ECM}(h) = \frac{1}{T} \sum_t^T (Y_{t+h} - y_t^f(h))^2$$

Las últimas filas de la Tabla 2 nos muestran dos hechos estilizados que nos parecen interesantes de destacar y que se verifican en las dos muestras consideradas. En primer lugar se observa que las proyecciones que presumiblemente se construyen con mayor información presentan un menor ECM estimado. Así, se verifica que las proyecciones de Bloomberg son más precisas que las de Consensus, las de Consensus más precisas que las de la EEE, y las proyecciones de la EEE más precisas que las del promedio de modelos ESARIMA. En segundo lugar, también se destaca que para todas las proyecciones el ECM crece con el horizonte de proyección, lo que, al igual que lo destacado en el primer punto anterior, también es coherente con un perfil de pronósticos óptimo bajo pérdida cuadrática.

En la Tabla 3 se muestra el porcentaje de veces en que una proyección arroja el menor error. La suma puede ser mayor a 100% toda vez que algunas proyecciones sean iguales. Se observa que ninguna lo hace mejor el 100% de las veces. De hecho, el mayor valor alcanzado es cercano al 80% para el caso de la EEE proyectando a doce meses y compitiendo con el promedio ESARIMA. En definitiva, no existe una dominancia uniforme de un pronóstico sobre otro sino que diferencias en el promedio de sus precisiones predictivas.

Otra característica importante del conjunto de pronósticos es la alta correlación entre ellos, al menos en los horizontes más cortos. En la Tabla 4 se presenta la matriz de correlaciones de los pronósticos para los dos horizontes predictivos en que existen más de dos proyecciones: $h = 1$ y 3 . Para los casos en que el horizonte predictivo es igual a $h = 2, 4, 5, 6$ y 12 , las correlaciones caen mientras mayor es el horizonte, pasando de $0,94$ ($h = 2$) a $0,01$ ($h = 12$).

Tabla 1: Períodos muestrales sin observaciones faltantes (*).

	ESARIMA	EEE	Consensus Economics	Bloomberg
h=1	Ene-2001 – Dic-2009 (108)	Ene-2001 – Dic-2009 (108)	Abr-2001 – Dic-2009 (105)	Dic-2003 – Dic-2009 (73)
h=2	Ene-2001 – Dic-2009 (108)	-	Abr-2001 – Dic-2009 (105)	-
h=3	Ene-2001 – Dic-2009 (108)	Nov-2001 – Dic-2009 (98)	Abr-2001 – Dic-2009 (105)	-
h=4	Ene-2001 – Dic-2009 (108)	-	Abr-2001 – Dic-2009 (105)	-
h=5	Ene-2001 – Dic-2009 (108)	-	Abr-2001 – Dic-2009 (105)	-
h=6	Ene-2001 – Dic-2009 (108)	-	Abr-2001 – Dic-2009 (105)	-
h=12	Ene-2001 – Dic-2009 (108)	Sep-2001 – Dic-2009 (100)	-	-

(*) Entre paréntesis se presenta el número de observaciones. Fuente: Elaboración propia.

Tabla 2: Estadísticos típicos de todas las series (*).

	Muestra completa (Tabla 1)					Muestra: Dic 2003 – Dic 2009 (73 datos)				
	Original	ESARIMA	EEE	C. Economics	Bloomberg	Original	ESARIMA	EEE	C. Economics	Bloomberg
Promedio	3.45	3.48	3.43	3.42	3.63	3.66	3.67	3.62	3.62	3.63
Desv. Est.	2.51	2.53	2.46	2.49	2.94	2.99	3.01	2.94	2.93	2.94
Máximo	9.90	9.96	9.54	9.50	9.50	9.90	9.96	9.54	9.50	9.50
Mínimo	-2.30	-2.10	-1.81	-1.90	-1.90	-2.30	-2.10	-1.81	-1.90	-1.90
Coef. Var.	0.72	0.72	0.71	0.72	0.81	0.81	0.82	0.81	0.81	0.81
RECM h=1	-	0.44	0.32	0.30	0.24	-	0.45	0.34	0.32	0.25
RECM h=2	-	0.81	-	0.64	-	-	0.88	-	0.71	-
RECM h=3	-	1.12	0.99	0.80	-	-	1.26	1.05	0.89	-
RECM h=4	-	1.43	-	0.96	-	-	1.65	-	1.10	-
RECM h=5	-	1.76	-	1.14	-	-	2.05	-	1.32	-
RECM h=6	-	2.08	-	1.30	-	-	2.44	-	1.53	-
RECM h=12	-	3.48	2.06	-	-	-	4.15	2.34	-	-

(*) Promedio, desv. est., máximo y mínimo de la serie un mes adelante, medida en puntos porcentuales. RECM $h=j$ significa raíz del error cuadrático medio j meses adelante. La 'Muestra completa' varía según lo reportado en la Tabla 1. Para la serie original se considera una muestra desde Enero del 2001 hasta Diciembre del 2009. Fuente: Elaboración propia.

Tabla 3: Porcentaje de veces superior (*).

	Muestra: Diciembre 2003 – Diciembre 2009 (73 datos)			
	ESARIMA	EEE	Cons. Frct.	Bloomberg
h=1	21.9 %	12.3 %	24.7 %	46.6 %
h=2	42.5 %	-	57.5 %	-
h=3	19.2 %	19.2 %	37.0 %	-
h=4	42.5 %	-	57.5 %	-
h=5	43.8 %	-	56.2 %	-
h=6	37.0 %	-	63.0 %	-
h=12	19.2 %	80.8 %	-	-

(*) Fuente: Elaboración propia.

Tabla 4: Matriz de correlaciones (*).

	Proyecciones							
	h=1				h=3			
	ESARIMA	EEE	Cons. Frct.	Bloomberg	ESARIMA	EEE	Cons. Frct.	Bloomberg
ESARIMA	1.0000	-	-	-	1.0000	-	-	-
EEE	0.9963	1.0000	-	-	0.9730	1.0000	-	-
Cons. Frct.	0.9957	0.9995	1.0000	-	0.8444	0.8934	1.0000	-
Bloomberg	0.9942	0.9985	0.9991	1.0000	-	-	-	-

4. Un Perfil de Proyecciones Optimo

Un resultado tradicional en la literatura predictiva es aquel que muestra que el pronóstico óptimo bajo pérdida cuadrática corresponde a la esperanza condicional de la variable a predecir, con respecto al conjunto de información disponible. Esto se muestra en el artículo de Elliot y Timmermann (2008), por citar un ejemplo. Este resultado conlleva a que el error de predicción óptimo bajo pérdida cuadrática cumpla las propiedades de insesgamiento y de ortogonalidad con respecto a cualquier variable presente en el conjunto de información. Una derivación de algunas de estas propiedades se puede encontrar en Elliot y Timmermann (2008), Bentancor y Pincheira (2010) y en Capistrán y López-Moctezuma (2010).

Más allá de las propiedades de sesgo y de ortogonalidad que caracterizan al error de predicción óptimo bajo pérdida cuadrática, en este artículo estamos interesados en dos propiedades particulares que restringen

de una manera particular a un perfil de pronósticos contruidos en distintos momentos del tiempo para la misma variable Y_{t+h} . La idea principal es revisar la consistencia interna de este perfil de proyecciones bajo el supuesto de que han sido contruidos minimizando una función de pérdida cuadrática.

En las próximas líneas seguiremos de cerca a Patton y Timmermann (2010) para derivar las propiedades del perfil de pronósticos que estamos interesados en evaluar.

Consideremos el siguiente proceso:

$$Y = \{Y_t\}_{t=0}^{\infty}$$

que supondremos adaptado a una filtración denominada $\mathcal{F} = \{\mathcal{F}_t\}_{t=0}^{\infty}$. Recordemos que al ser \mathcal{F} una filtración se cumple que:

$$\mathcal{F}_t \subseteq \mathcal{F}_{t'}, t' > t$$

Cada uno de los elementos de la filtración \mathcal{F}_t constituye el conjunto de información disponible en el instante t . Supondremos además que la siguiente esperanza condicional existe:

$$\mathbb{E}[Y_t | \mathcal{F}_{t-h}] \text{ para todo } t, h = 0, 1, 2, \dots; \text{ tal que } t - h \geq 0$$

Tal como hemos señalado anteriormente, el pronóstico óptimo bajo pérdida cuadrática corresponde a la esperanza condicional de la variable a predecir con respecto al conjunto de información utilizado en la predicción:

$$y_{t-h}^f(h)^* \equiv \mathbb{E}[Y_t | \mathcal{F}_{t-h}]$$

Siempre será posible descomponer la serie de nuestro interés de la siguiente manera:

$$Y_t = y_{t-h}^f(h)^* + e_{t-h}(h)^*$$

en que $e_{t-h}(h)^*$ denota al error de proyección óptimo definido como:

$$e_{t-h}(h)^* \equiv Y_t - y_{t-h}^f(h)^*$$

El ECM se define consecuentemente como:

$$ECM \equiv \mathbb{E}(e_{t-h}(h)^*)^2$$

mientras que el PCM queda simplemente definido como:

$$PCM \equiv \mathbb{E}(y_{t-h}^f(h)^*)^2$$

A continuación mostramos las propiedades que un perfil óptimo de proyecciones debe cumplir en términos de ECM y PCM.

4.1. Perfil Óptimo de Errores Cuadráticos Medios

Si nuestro objetivo es predecir Y_t y nuestro conjunto de información está dado por \mathcal{F}_{t-h_1} , $h_1 > 0$, $t - h_1 > 0$, entonces se cumple que

$$\mathbb{E} \left[(Y_t - y_{t-h_1}^f(h_1)^*)^2 \right] \leq \mathbb{E} [(Y_t - u)^2] \text{ para todo } u \in \mathcal{F}_{t-h_1}.$$

en particular si tomamos

$$\bar{u} = \mathbb{E}[Y_t | \mathcal{F}_{t-h_2}]; \quad h_1 < h_2 \text{ tal que } 0 \leq t - h_2$$

se tiene

$$\bar{u} \in \mathcal{F}_{t-h_1}$$

porque

$$0 \leq t - h_2 < t - h_1$$

y por lo tanto

$$\bar{u} = \mathbb{E}[Y_t | \mathcal{F}_{t-h_2}] \in \mathcal{F}_{t-h_2} \subseteq \mathcal{F}_{t-h_1}$$

luego

$$\mathbb{E} \left[(Y_t - y_{t-h_1}^f(h_1)^*)^2 \right] \leq \mathbb{E} \left[(Y_t - y_{t-h_2}^f(h_2)^*)^2 \right] \text{ para todo } h_1 < h_2 \text{ tal que } 0 \leq t - h_2.$$

lo que dice simplemente que el ECM del pronóstico óptimo bajo pérdida cuadrática decrece débilmente (no crece) en la medida que el pronóstico se construye con mayor información. Esto es cada vez más cerca del instante t en el que se observa la variable a predecir.

Si ahora nuestro objetivo es predecir Y_{t+h} y nuestro conjunto de información está dado por \mathcal{F}_t , $h > 0, t \geq 0$, entonces se cumple que

$$\mathbb{E} \left[(Y_{t+h} - y_t^f(h)^*)^2 \right] \leq \mathbb{E} \left[(Y_{t+h} - u)^2 \right] \text{ para todo } u \in \mathcal{F}_t$$

en particular si tomamos

$$\bar{u} = \mathbb{E}[Y_{t+h} | \mathcal{F}_{t-k}]; \quad 0 < k \text{ tal que } 0 \leq t - k$$

se tiene

$$\bar{u} \in \mathcal{F}_t$$

porque

$$0 \leq t - k < t$$

y por lo tanto

$$\bar{u} = \mathbb{E}[Y_{t+h} | \mathcal{F}_{t-k}] \in \mathcal{F}_{t-k} \subseteq \mathcal{F}_t$$

luego

$$\mathbb{E} \left[(Y_{t+h} - y_t^f(h)^*)^2 \right] \leq \mathbb{E} \left[(Y_{t+h} - y_{t-k}^f(h+k)^*)^2 \right] \text{ para todo } 0 < k \text{ tal que } 0 \leq t - k.$$

lo que dice simplemente que el ECM del pronóstico óptimo bajo pérdida cuadrática crece débilmente (no decrece) con el horizonte de proyección.

4.2. Perfil Óptimo de Pronósticos Cuadráticos Medios

Dado que bajo optimalidad se cumple que

$$\begin{aligned} \mathbb{E}[e_{t-h}(h)^*] &= \mathbb{E}[\mathbb{E}[e_{t-h}(h)^* | \mathcal{F}_{t-h}]] = 0 \\ \text{Cov}[y_{t-h}^f(h)^*, e_{t-h}(h)^*] &= \mathbb{E}[y_{t-h}^f(h)^* e_{t-h}(h)^*] = \mathbb{E}[\mathbb{E}[y_{t-h}^f(h)^* e_{t-h}(h)^* | \mathcal{F}_{t-h}]] \\ &= \mathbb{E}[y_{t-h}^f(h)^* \mathbb{E}[e_{t-h}(h)^* | \mathcal{F}_{t-h}]] = 0 \end{aligned}$$

entonces

$$\begin{aligned} \mathbb{V}[Y_t] &= \mathbb{V}[y_{t-h}^f(h)^* + e_{t-h}(h)^*] \\ &= \mathbb{V}[y_{t-h}^f(h)^*] + \mathbb{V}[e_{t-h}(h)^*] \end{aligned}$$

lo que se puede reescribir como:

$$\mathbb{V}[y_{t-h}^f(h)^*] = \mathbb{V}[Y_t] - \mathbb{V}[e_{t-h}(h)^*]$$

que a su vez se puede escribir como

$$\mathbb{E} \left[y_{t-h}^f(h)^* \right]^2 - \left[\mathbb{E} \left(y_{t-h}^f(h)^* \right) \right]^2 = \mathbb{E} [Y_t]^2 - [\mathbb{E} (Y_t)]^2 - \mathbb{E} [e_{t-h}(h)^*]^2 + [\mathbb{E} (e_{t-h}(h)^*)]^2 \quad (1)$$

Pero recordemos que

$$\mathbb{E} (e_{t-h}(h)^*) = 0$$

por lo que

$$\mathbb{E} (Y_t) = \mathbb{E} \left(y_{t-h}^f(h)^* + e_{t-h}(h)^* \right) = \mathbb{E} \left(y_{t-h}^f(h)^* \right)$$

Así (1) se puede expresar equivalentemente como

$$\mathbb{E} \left[y_{t-h}^f(h)^* \right]^2 - [\mathbb{E} (Y_t)]^2 = \mathbb{E} [Y_t]^2 - [\mathbb{E} (Y_t)]^2 - \mathbb{E} [e_{t-h}(h)^*]^2$$

o bien como

$$\mathbb{E} \left[y_{t-h}^f(h)^* \right]^2 = \mathbb{E} [Y_t]^2 - \mathbb{E} [e_{t-h}(h)^*]^2$$

Esta expresión indica que el Pronóstico Cuadrático Medio (PCM) del pronóstico óptimo bajo pérdida cuadrática presenta un comportamiento opuesto al del ECM, esto es, crece débilmente (no decrece) en la medida que el pronóstico se construye con mayor información, es decir, cada vez más cerca del instante t en el que se observa la variable a predecir.

Si ahora nuestro objetivo es predecir Y_{t+h} y nuestro conjunto de información está dado por \mathcal{F}_t , $h > 0$, $t \geq 0$, entonces se cumple que

$$\begin{aligned} \mathbb{V} [Y_{t+h}] &= \mathbb{V} \left[y_t^f(h)^* + e_t(h)^* \right] \\ &= \mathbb{V} \left[y_t^f(h)^* \right] + \mathbb{V} [e_t(h)^*] \end{aligned}$$

dada la ortogonalidad de $y_t^f(h)^*$ y $e_t(h)^*$.

Esta expresión se puede reescribir como

$$\mathbb{V} \left[y_t^f(h)^* \right] = \mathbb{V} [Y_t] - \mathbb{V} [e_t(h)^*]$$

que a su vez se puede plantear como

$$\mathbb{E} \left[y_t^f(h)^* \right]^2 = \mathbb{E} [Y_t]^2 - \mathbb{E} [e_t(h)^*]^2$$

lo que dice simplemente que el PCM del pronóstico óptimo bajo pérdida cuadrática decrece débilmente (no crece) con el horizonte de proyección.

En esta subsección hemos resumido cuatro propiedades de optimalidad que han sido recientemente propuestas en Patton y Timmermann (2010). En la siguiente subsección exploramos preliminarmente si estas condiciones de optimalidad son satisfechas o no para el perfil de proyecciones de inflación en Chile públicamente disponible. En secciones posteriores estudiamos estas condiciones de optimalidad en un contexto formal de inferencia estadística.

4.3. Estadística Descriptiva del Perfil de Proyecciones de Inflación en Chile

En la Tabla 5 se presentan los estimadores puntuales del ECM para todas las proyecciones de inflación que consideramos en este artículo. Los resultados se presentan para la muestra total disponible, así como para la submuestra que comprende el período Diciembre del 2003 a Diciembre del 2009. Al observar cada columna de la Tabla 5 se encuentra que efectivamente el estimador puntual del ECM crece con el horizonte de proyección para todos los perfiles que contienen pronósticos a más de un horizonte (ESARIMA, EEE y Consensus), tanto para la muestra completa como la submuestra. Similarmente, y al observar cada fila de la Tabla 5 desde izquierda a derecha, se observa un perfil decreciente en los estimadores del ECM para cada horizonte de proyección. Esto ocurre para los dos periodos muestrales considerados. Este resultado es consistente con las propiedades de optimalidad anteriormente demostradas, puesto que muestra que las proyecciones construidas con menor información presentan un estimador puntual del ECM mayor que el de proyecciones construidas con mayor información.

Tabla 5: Errores Cuadrático Medio (*).

ECM	Muestra completa (Tabla 1)				Muestra: Dic-2003 – Dic-2009 (73 datos)			
	ESARIMA	EEE	Cons. Frct.	Bloomberg	ESARIMA	EEE	Cons. Frct.	Bloomberg
h=1	0.19	0.10	0.09	0.06	0.20	0.11	0.10	0.06
h=2	0.66	-	0.41	-	0.78	-	0.51	-
h=3	1.25	0.99	0.64	-	1.58	1.10	0.79	-
h=4	2.05	-	0.93	-	2.73	-	1.21	-
h=5	3.09	-	1.29	-	4.21	-	1.74	-
h=6	4.31	-	1.70	-	5.93	-	2.32	-
h=12	12.12	4.25	-	-	17.24	5.48	-	-

(*) Fuente: Elaboración propia.

En la Tabla 6 se presentan los estimadores puntuales de los PCM para todos los perfiles de proyección en las mismas dos muestras consideradas en la Tabla 5. Al observar cada columna de la Tabla 6 se encuentra que el estimador puntual del PCM decrece con el horizonte de proyección para casi todos los perfiles considerados (EEE, Consensus y Bloomberg), tanto para la muestra completa como la submuestra. La única excepción se detecta para el caso de las proyecciones ESARIMA que presentan un perfil creciente en los seis primeros horizontes de proyección para luego decaer en el horizonte doce. Este comportamiento creciente en los primeros seis meses no es compatible con un perfil de proyecciones óptimo de acuerdo a lo demostrado en secciones anteriores.

Si ahora observamos cada fila de la Tabla 6 encontramos mayoritariamente un ordenamiento incompatible con propiedades de optimalidad. Por ejemplo, para el primer horizonte se observa que el estimador puntual del PCM de la proyección ESARIMA es mayor que el PCM estimado para la EEE y Consensus en ambos periodos muestrales considerados. Para este mismo horizonte de proyección, en el periodo muestral que va desde Diciembre del 2003 a Diciembre del 2009 es posible detectar aún más desviaciones de optimalidad. En efecto, a la anomalía ya mencionada en el punto anterior, ahora se agrega que el estimador puntual del PCM de los modelos ESARIMA es también mayor que el de Bloomberg, y que este estimador para la EEE también es mayor que el de Consensus, lo que es incoherente con un perfil de proyecciones óptimo bajo pérdida cuadrática. El análisis del resto de los horizontes de proyección no hace más que confirmar las incompatibilidades con optimalidad mencionadas anteriormente. De hecho, más allá del primer horizonte, no existe un par de proyecciones que presente estimadores puntuales de PCM compatibles con optimalidad.

Tabla 6: Proyecciones Cuadrático Medias (*).

PCM	Muestra completa (Tabla 1)				Muestra: Dic-2003 – Dic-2009 (73 datos)			
	ESARIMA	EEE	Cons. Frct.	Bloomberg	ESARIMA	EEE	Cons. Frct.	Bloomberg
h=1	18.485	17.850	17.858	21.770	22.486	21.641	21.634	21.770
h=2	18.697	-	17.187	-	22.585	-	20.694	-
h=3	18.908	17.243	16.478	-	22.753	20.004	19.622	-
h=4	19.043	-	15.655	-	22.883	-	18.452	-
h=5	19.115	-	14.712	-	22.944	-	17.169	-
h=6	19.208	-	13.835	-	23.100	-	15.997	-
h=12	18.522	10.002	-	-	22.457	10.296	-	-

(*) Fuente: Elaboración propia.

En la próxima sección veremos si los resultados mostrados en las tablas 5 y 6 contienen suficiente información como para rechazar la hipótesis nula de optimalidad.

5. Inferencia Estadística Acerca del Perfil de Proyecciones de Inflación en Chile

En esta sección deseamos contrastar diversas hipótesis nulas que son coherentes con la existencia de perfiles de proyección óptimos bajo pérdida cuadrática.

La primera hipótesis nula de interés, que denotamos H_{01}^h , plantea para cada horizonte predictivo, que aquel pronóstico que se construye en un instante más cercano a la publicación de la variable de interés debería presentar un ECM menor o igual al de aquellos pronósticos que se construyen en instantes más alejados de la publicación del dato de inflación relevante. Así tenemos que para cualquier horizonte predictivo h :

$$H_{01}^h : ECM_{Bloomberg}^h \leq ECM_{Consensus}^h \leq ECM_{EEE}^h \leq ECM_{ESARIMA}^h$$

Nuestra segunda hipótesis nula de interés, que denotamos, H_{02}^h plantea, para cada horizonte predictivo, que aquel pronóstico que se construye en un instante más cercano a la publicación de la variable de interés debería presentar un PCM mayor o igual al de aquellos pronósticos que se construyen en instantes más alejados de la publicación del dato de inflación relevante. Así tenemos que para cualquier horizonte predictivo h :

$$H_{02}^h : PCM_{Bloomberg}^h \geq PCM_{Consensus}^h \geq PCM_{EEE}^h \geq PCM_{ESARIMA}^h$$

La tercera hipótesis nula que nos interesa estudiar se refiere a la consistencia temporal de cada perfil de proyecciones en términos de ECM. Esto significa que tanto para las proyecciones de Bloomberg, como para las de Consensus, así como para las de la EEE y las ESARIMA, se debería verificar que las proyecciones a horizontes cortos debiesen tener un ECM menor o igual al de las proyecciones de más largo plazo. Así, para cada fuente de proyección $j \in \{Bloomberg, Consensus, EEE, ESARIMA\}$ la siguiente hipótesis nula H_{03}^j debería cumplirse:

$$H_{03}^j : ECM(h = 1) \leq ECM(h = 2) \leq \dots \leq ECM(h = 6) \leq \dots \leq ECM(h = 12)$$

en que $h = 1, \dots, 12$ denotan los horizontes de proyección.

Finalmente la cuarta hipótesis a evaluar es similar a la anterior pero ahora se refiere a la consistencia temporal de cada perfil de proyecciones en términos de PCM. Esto significa que tanto para las proyecciones de Bloomberg, como para las de Consensus, así como para las de la EEE y las ESARIMA, se debería verificar que las proyecciones a horizontes cortos debiesen tener un PCM mayor o igual al de las proyecciones de más largo plazo. Así, para cada fuente de proyección $j \in \{Bloomberg, Consensus, EEE, ESARIMA\}$ la siguiente hipótesis nula H_{04}^j debería cumplirse:

$$H_{04}^j : PCM(h = 1) \geq PCM(h = 2) \geq \dots \geq PCM(h = 6) \geq \dots \geq PCM(h = 12)$$

en que $h = 1, \dots, 12$ denotan los horizontes de proyección.

Las cuatro hipótesis nulas formuladas anteriormente son similares. En particular, todas ellas corresponden a un conjunto de desigualdades entre momentos de diversas variables aleatorias. Notamos también que en muchas ocasiones nuestras hipótesis de interés corresponden a hipótesis univariadas, que involucran solo una desigualdad entre dos momentos poblacionales. Esto ocurre para las hipótesis $H_{01}^2, H_{02}^2, H_{01}^4, H_{02}^4, H_{01}^5, H_{02}^5, H_{01}^6, H_{02}^6$ y H_{01}^{12}, H_{02}^{12} . En la siguiente subsección analizaremos la hipótesis de optimalidad para el perfil de proyecciones en aquellos horizontes predictivos que solo involucran la comparación de dos proyecciones.

5.1. Implementación de Tests Univariados

Para el caso de aquellos horizontes predictivos en los que solo se tiene disponible un par de proyecciones distintas, una forma fácil de proceder consiste en utilizar algún test de igualdad de capacidad predictiva pero a una sola cola. Alternativas naturales de estos tests son los atribuidos a Diebold y Mariano (1995) y West (1996) o bien la versión incondicional del test de Giacomini y White (2006), que denotaremos por GW. Formalmente, para estos casos utilizaremos el siguiente estadístico de GW:

$$t_{n(h)} = \frac{\overline{\Delta L_h}}{\widehat{\sigma}_{n(h)}/\sqrt{n(h)}},$$

con:

$$\overline{\Delta L_h} = \frac{1}{n(h)} \sum_{t=1}^{n(h)} \Delta L_{th},$$

donde ΔL_h es el diferencial de la función de pérdida entre dos series de pronósticos, $n(h)$ es el número de observaciones (proyecciones) a un horizonte h , y $\widehat{\sigma}_{n(h)}$ es un estimador HAC de la desviación estándar de $\overline{\Delta L_h} \sqrt{n(h)}$. Giacomini y White (2006) muestran que este estadístico tiene una distribución asintótica normal estándar. Cuando estamos interesados en comparaciones de ECM, entonces el diferencial de funciones de pérdida entre la serie de pronósticos A y B luce como sigue:

$$\Delta L_{th} = e_{A,t}^2(h) - e_{B,t}^2(h)$$

Mientras en los casos en los que estamos interesados en comparaciones de PCM el diferencial de funciones de pérdida entre la serie de pronósticos A y B luce como sigue:

$$\Delta L_{th} = (y_{A,t}^f(h))^2 - (y_{B,t}^f(h))^2$$

En síntesis, la evaluación de un perfil óptimo de proyecciones de inflación se puede resolver satisfactoriamente a través de la aplicación de un test de GW a una sola cola para contrastar las siguientes hipótesis nulas:

$$\begin{aligned} H_{01}^2 & : ECM_{Consensus}(h=2) - ECM_{ESARIMA}(h=2) \leq 0 \\ H_{01}^4 & : ECM_{Consensus}(h=4) - ECM_{ESARIMA}(h=4) \leq 0 \\ H_{01}^5 & : ECM_{Consensus}(h=5) - ECM_{ESARIMA}(h=5) \leq 0 \\ H_{01}^6 & : ECM_{Consensus}(h=6) - ECM_{ESARIMA}(h=6) \leq 0 \\ H_{01}^{12} & : ECM_{ESARIMA}(h=12) - ECM_{EEE}(h=12) \leq 0 \end{aligned}$$

y

$$\begin{aligned} H_{02}^2 & : PCM_{Consensus}(h=2) - PCM_{ESARIMA}(h=2) \geq 0 \\ H_{02}^4 & : PCM_{Consensus}(h=4) - PCM_{ESARIMA}(h=4) \geq 0 \\ H_{02}^5 & : PCM_{Consensus}(h=5) - PCM_{ESARIMA}(h=5) \geq 0 \\ H_{02}^6 & : PCM_{Consensus}(h=6) - PCM_{ESARIMA}(h=6) \geq 0 \\ H_{02}^{12} & : PCM_{ESARIMA}(h=12) - PCM_{EEE}(h=12) \geq 0 \end{aligned}$$

Para la evaluación en ECM, el test rechazará la hipótesis nula si

$$t_{n(h)} = \frac{\overline{\Delta L_h}}{\widehat{\sigma}_{n(h)}/\sqrt{n(h)}} > c_1 > 0$$

mientras que para la evaluación de PCM, el test rechazará si

$$t_{n(h)} = \frac{\overline{\Delta L_h}}{\widehat{\sigma}_{n(h)}/\sqrt{n(h)}} < c_2 < 0$$

en que c_1 y c_2 son valores críticos que provienen de una distribución normal estándar.

En las Tablas 7 y 8 se muestran los resultados del test de GW a una cola para los ECM y PCM respectivamente. Estas tablas contienen los resultados de los tests para todas las posibles comparaciones bilaterales entre proyecciones. No obstante, recordemos que en esta parte del artículo hemos decidido centrarnos solo en aquellas hipótesis univariadas, es decir, no conjuntas.

Tabla 7: Test de GW para ECM (*).

	Muestra completa (Tabla 1)					
	Bloomberg y Cons. Frct.	Bloomberg y EEE	Bloomberg y ESARIMA	Cons. Frct. y EEE	Cons. Frct. y ESARIMA	EEE y ESARIMA
h=1	-0.038 (0.010)	-0.051 (0.010)	-0.143 (0.044)	-0.013 (0.006)	-0.104 (0.039)	-0.091 (0.037)
h=2	- -	- -	- -	- -	-0.269 (0310)	- -
h=3	- -	- -	- -	-0.303 (0.328)	-0.786 (0.488)	-0.482 (0.416)
h=4	- -	- -	- -	- -	-1.519 (0.776)	- -
h=5	- -	- -	- -	- -	-2.465 (1.145)	- -
h=6	- -	- -	- -	- -	-3.602 (1.694)	- -
h=12	- -	- -	- -	- -	- -	-11.759 (6.684)

(*) Entre paréntesis se presenta la desviación estándar. (†) Significativo al 10%.

(††) Significativo al 5%. Fuente: Elaboración propia.

La Tabla 7 nos indica que las hipótesis nulas $H_{01}^2, H_{01}^4, H_{01}^5, H_{01}^6$ y H_{01}^{12} no pueden ser rechazadas a ningún nivel de significancia usual (10%, 5% o 1%). De hecho, esta tabla nos indica que cualquier hipótesis nula bilateral planteando superioridad en términos de ECM por parte de los pronósticos construidos en un instante más cercano a la publicación de la variable de interés, no podría ser rechazada a niveles de significancia usuales. Aún más, todos los signos de los estadísticos apuntan en la dirección de establecer un perfil de estimaciones puntuales de ECM coherentes con un perfil de proyecciones óptimas a cualquier horizonte predictivo.

La Tabla 8 nos muestra un panorama algo distinto. Si bien esta tabla indica que las hipótesis nulas $H_{02}^2, H_{02}^4, H_{02}^5, H_{02}^6$ tampoco pueden ser rechazadas a ningún nivel de significancia usual (10%, 5% o 1%), la hipótesis H_{02}^{12} efectivamente es rechazada a un nivel de significancia del 5%. Más aún, esta tabla nos indica que en varias ocasiones una hipótesis nula bilateral, planteando que los pronósticos que se construyen en un instante más cercano a la publicación de la variable de interés debería presentar un PCM mayor o igual al

de aquellos pronósticos que se construyen en instantes más alejados de la publicación oficial, sería rechazada a niveles de significancia usuales. De hecho, varios signos de los estadísticos apuntan en la dirección de establecer un perfil de estimaciones puntuales de PCM que no son coherentes con un perfil de proyecciones óptimas bajo pérdida cuadrática.

Tabla 8: Test de GW para PCM(*).

Muestra completa (Tabla 1)						
	Bloomberg y Cons. Frct.	Bloomberg y EEE	Bloomberg y ESARIMA	Cons. Frct. y EEE	Cons. Frct. y ESARIMA	EEE y ESARIMA
h=1	0.135 (0.222)	0.128 (0.292)	-0.716 [†] (0.539)	-0.006 (0.135)	-0.851 ^{††} (0.451)	-0.845 ^{††} (0.378)
h=2	- -	- -	- -	- -	-1.890 (1.576)	- -
h=3	- -	- -	- -	-0.382 (2.101)	-3.131 (2.790)	-2.748 ^{††} (1.214)
h=4	- -	- -	- -	- -	-4.431 (3.966)	- -
h=5	- -	- -	- -	- -	-5.775 (5.044)	- -
h=6	- -	- -	- -	- -	-7.103 (5.859)	- -
h=12	- -	- -	- -	- -	- -	-12.161 ^{††} (6.666)

(*) Entre paréntesis se presenta la desviación estándar. (†) Significativo al 10 %.

(††) Significativo al 5 %. Fuente: Elaboración propia.

Hasta ahora hemos dado una respuesta satisfactoria a la evaluación de propiedades de optimalidad para el perfil de proyecciones en aquellos horizontes en los que solo tenemos dos proyecciones. En la siguiente subsección analizaremos lo que ocurre en aquellos horizontes en los que hay más de un par de proyecciones disponibles.

5.2. Implementación de Tests Multivariados o Conjuntos

Las tablas 6 y 7 nos muestran que para el primer y el tercer horizonte de proyección las hipótesis nulas $H_{01}^1, H_{01}^3, H_{02}^1$ y H_{02}^3 son hipótesis que relacionan tres o más series de pronósticos, lo que conlleva la necesidad de trabajar con un test conjunto.

Como hemos señalado en la introducción, el tema de fondo que subyace a toda hipótesis nula multivariada se relaciona con la necesidad y dificultad de controlar correctamente por el error de tipo I de la prueba estadística. Antes de pasar de lleno a la construcción de un test conjunto para contrastar nuestras hipótesis multivariadas, estudiemos que ocurre con ellas si aplicamos las siguientes reglas simples:

1. Rechazar la nula a un nivel de significancia α , si cualquiera de las comparaciones bilaterales contiguas es rechazada a un nivel de significancia α usando el test de GW a una sola cola.
2. Rechazar la nula a un nivel de significancia α , si cualquiera de las comparaciones bilaterales contiguas es rechazada a un nivel de significancia α/N usando el test de GW a una sola cola, en que N representa al número de comparaciones bilaterales contiguas. Esta regla corresponde al uso de las cotas de Bonferroni.

Por ejemplo, analicemos el caso del ECM de las proyecciones 3 meses hacia adelante. En este caso la hipótesis nula es denotada por H_{01}^3 y definida como sigue:

$$H_{01}^3 : ECM_{Consensus}^3 \leq ECM_{EEE}^3 \leq ECM_{ESARIMA}^3$$

En estricto rigor esta hipótesis nula plantea tres desigualdades:

$$\begin{aligned} ECM_{Consensus}^3 &\leq ECM_{EEE}^3 \\ ECM_{Consensus}^3 &\leq ECM_{ESARIMA}^3 \\ ECM_{EEE}^3 &\leq ECM_{ESARIMA}^3 \end{aligned}$$

Pero notemos que el conjunto paramétrico definido por H_{01}^3 se puede expresar equivalentemente solo en términos de las siguientes dos desigualdades contiguas:

$$H_{01}^3 : ECM_{Consensus}^3 \leq ECM_{EEE}^3 \text{ y } ECM_{EEE}^3 \leq ECM_{ESARIMA}^3$$

Esta observación nos permite evaluar las hipótesis nulas anidadas en términos de un conjunto menor de desigualdades contiguas, lo que haremos a continuación:

5.2.1. Evaluaciones Conjuntas: Una Primera Aproximación

En esta sección deseamos evaluar las siguientes hipótesis nulas:

$$\begin{aligned} H_{01}^1 &: ECM_{Bloomberg}^1 \leq ECM_{Consensus}^1; ECM_{Consensus}^1 \leq ECM_{EEE}^1 \text{ y } ECM_{EEE}^1 \leq ECM_{ESARIMA}^1 \\ H_{01}^3 &: ECM_{Consensus}^3 \leq ECM_{EEE}^3 \text{ y } ECM_{EEE}^3 \leq ECM_{ESARIMA}^3 \\ H_{02}^1 &: PCM_{Bloomberg}^1 \geq PCM_{Consensus}^1; PCM_{Consensus}^1 \geq PCM_{EEE}^1 \text{ y } PCM_{EEE}^1 \geq PCM_{ESARIMA}^1 \\ H_{02}^3 &: PCM_{Consensus}^3 \geq PCM_{EEE}^3 \text{ y } PCM_{EEE}^3 \geq PCM_{ESARIMA}^3 \end{aligned}$$

Definamos un nivel de significancia del 10% y evaluemos estas 4 hipótesis de acuerdo a las reglas 1 y 2. Observamos que las hipótesis H_{01}^1 y H_{01}^3 no son rechazadas de acuerdo a ninguna de las dos reglas, puesto que ninguna de las relaciones bilaterales es rechazada ni siquiera al 10% como puede apreciarse en la Tabla 7. Un escenario opuesto encontramos cuando evaluamos las hipótesis H_{02}^1 y H_{02}^3 . No es difícil calcular los *p-value* del test de GW a una sola cola para evaluar las siguientes dos hipótesis bilaterales:

$$\begin{aligned} PCM_{EEE}^1 - PCM_{ESARIMA}^1 &\geq 0 \\ PCM_{EEE}^3 - PCM_{ESARIMA}^3 &\geq 0 \end{aligned}$$

Estas hipótesis bilaterales se rechazan a un nivel de significancia del 5%. De hecho los respectivos *p-value* corresponden a 0.013 y 0.012 respectivamente. Como consecuencia, tanto la regla 1 como la 2 rechazan las hipótesis nulas H_{02}^1 y H_{02}^3 a un nivel de significancia del 10%, e incluso del 5% también.

Las siguientes tablas 9 y 10 nos muestran los resultados de los test de GW, pero esta vez practicado sobre cada conjunto de proyecciones. Así evaluamos si el ECM de las proyecciones ESARIMA, por ejemplo, presenta un comportamiento creciente con el horizonte de proyección. La tabla 9 en particular se centra en evaluar el perfil creciente de los ECM para las proyecciones ESARIMA, de la EEE y Consensus. La tabla 10 evalúa el supuesto perfil decreciente de los PCM para los respectivos conjuntos de proyecciones.

Nuestro objetivo es evaluar de acuerdo a las reglas 1 y 2 las siguientes hipótesis:

$$\begin{aligned} H_{03}^{Consensus} &: ECM_{Consensus}^{h=1} \leq ECM_{Consensus}^{h=2} \leq \dots \leq ECM_{Consensus}^{h=6} \\ H_{03}^{EEE} &: ECM_{EEE}^{h=1} \leq ECM_{EEE}^{h=3} \leq ECM_{EEE}^{h=12} \\ H_{03}^{ESARIMA} &: ECM_{ESARIMA}^{h=1} \leq ECM_{ESARIMA}^{h=2} \leq \dots \leq ECM_{ESARIMA}^{h=6} \leq ECM_{ESARIMA}^{h=12} \end{aligned}$$

$$\begin{aligned}
H_{04}^{Consensus} & : PCM_{Consensus}^{h=1} \geq PCM_{Consensus}^{h=2} \geq \dots \geq PCM_{Consensus}^{h=6} \\
H_{04}^{EEE} & : PCM_{EEE}^{h=1} \geq PCM_{EEE}^{h=3} \geq PCM_{EEE}^{h=12} \\
H_{04}^{ESARIMA} & : PCM_{ESARIMA}^{h=1} \geq PCM_{ESARIMA}^{h=2} \geq \dots \geq PCM_{ESARIMA}^{h=6} \geq PCM_{ESARIMA}^{h=12}
\end{aligned}$$

en que $h = 1, \dots, 12$ denotan los horizontes de proyección.

Al igual que lo hecho anteriormente, aplicaremos las reglas 1 y 2 usando los test de GW bilaterales entre horizontes adyacentes. Así por ejemplo, rechazaremos la hipótesis nula $H_{03}^{Consensus}$ con un nivel de significancia del 10% si al menos uno de los test de GW para las siguientes hipótesis

$$\begin{aligned}
ECM_{Consensus}^{h=1} - ECM_{Consensus}^{h=2} & \leq 0 \\
ECM_{Consensus}^{h=2} - ECM_{Consensus}^{h=3} & \leq 0 \\
ECM_{Consensus}^{h=3} - ECM_{Consensus}^{h=4} & \leq 0 \\
ECM_{Consensus}^{h=4} - ECM_{Consensus}^{h=5} & \leq 0 \\
ECM_{Consensus}^{h=5} - ECM_{Consensus}^{h=6} & \leq 0
\end{aligned}$$

se rechaza a un nivel del 10% usando la regla 1, o bien a un nivel del 2% aplicando la regla 2.

Una rápida inspección a la tabla 9 nos indica que no existe ningún rechazo bilateral. De hecho todos los estadísticos de GW son negativos, lo que implica que no hay evidencia para rechazar la hipótesis nula de un perfil creciente de ECM para todos los conjuntos de proyecciones que poseen pronósticos a más de un horizonte.

Tabla 9: Resultados del test de GW entre horizontes para ECM (*).

Muestra: Dic-2003 - Dic-2009 (73 datos)

	ESARIMA	EEE	Cons. Frcst.
h=1 menos h=2	-0.574 (0.181)	-0.986 (0.275)	-0.408 (0.203)
h=2 menos h=3	-0.804 (0.305)	- -	-0.287 (0.195)
h=3 menos h=4	-1.146 (0.447)	-4.381 (1.476)	-0.413 (0.195)
h=4 menos h=5	-1.483 (0.621)	- -	-0.536 (0.207)
h=5 menos h=6	-1.716 (0.758)	- -	-0.579 (0.234)
h=6 menos h=12	-11.311 (6.138)	- -	- -

(*) Para los casos de EEE las comparaciones son: 1 - 3 y 6 - 12. (†) Significativo al 10%.

(††) Significativo al 5%. Entre paréntesis se presenta la desviación estándar. Fuente: Elaboración propia.

Similar resultado se obtiene cuando se analiza la tabla 10. Al no haber ni siquiera un rechazo al 10% usando el test de GW, no se puede rechazar la hipótesis nula de un perfil de PCM decreciente con el horizonte de predicción para ninguno de los conjuntos de pronósticos considerados: ESARIMA, EEE y Consensus.

Tabla 10: Resultados del test de GW entre horizontes para PCM (*).

Muestra: Dic-2003 - Dic-2009 (73 datos)

	ESARIMA	EEE	Cons. Frct.
h=1 menos h=2	-0.099 (1.017)	1.636 (1.923)	0.940 (0.627)
h=2 menos h=3	-0.167 (1.114)	- -	1.072 (0.590)
h=3 menos h=4	-0.130 (1.179)	9.708 (4.398)	1.169 (0.559)
h=4 menos h=5	-0.060 (1.248)	- -	1.283 (0.582)
h=5 menos h=6	-0.156 (1.257)	- -	1.171 (0.636)
h=6 menos h=12	0.643 (6.489)	- -	- -

(*) Para los casos de EEE las comparaciones son: 1 - 3 y 6 - 12. (†) Significativo al 10%.

(††) Significativo al 5%. Entre paréntesis se presenta la desviación estándar. Fuente: Elaboración propia.

En la siguiente sección veremos si todos estos resultados que hemos obtenidos aplicando las cuestionables reglas 1 y 2 sobreviven a la aplicación de un test conjunto más apropiado.

5.2.2. Un Test Conjunto de Superioridad Predictiva

En secciones precedentes señalamos que la contrastación de hipótesis conjuntas planteaba el desafío de diseñar tests que controlaran adecuadamente por el error de tipo I. En esta sección intentaremos hacer eso precisamente, es decir, diseñar un test de superioridad predictiva conjunto que controle adecuadamente por el error de tipo I. Más adelante, en otra sección, mostraremos con simulaciones de Monte Carlo por qué las reglas 1 y 2 definidas anteriormente pueden presentar severas deficiencias en el control del error de tipo I.

Consideremos la siguiente hipótesis nula genérica:

$$H_0^G : \mathbb{E}f(X_{1t}) \leq 0; \mathbb{E}f(X_{2t}) \leq 0; \dots; \mathbb{E}f(X_{lt}) \leq 0$$

contra la hipótesis alternativa genérica

$$H_A^G : \exists j \in \{1, 2, \dots, l\} \text{ tal que } \mathbb{E}f(X_{jt}) > 0$$

en que

$$X_t = (X_{1t}, \dots, X_{lt})'$$

representa a un vector de l variables aleatorias, y

$$f : \mathbb{R} \rightarrow \mathbb{R}$$

es una función continua. Denotaremos por Z al vector

$$Z_t = f(X_t)$$

Evidentemente nuestras hipótesis de interés son casos particulares de H_0^G . Por ejemplo, cuando la hipótesis de interés es

$$H_{01}^3 : ECM_{Consensus}^3 \leq ECM_{EEE}^3 \leq ECM_{ESARIMA}^3$$

basta tomar

$$\begin{aligned} X_{1t} &= e_{Consensus,t}^2(3) - e_{EEE,t}^2(3) \\ X_{2t} &= e_{Consensus,t}^2(3) - e_{ESARIMA,t}^2(3) \\ X_{3t} &= e_{EEE,t}^2(3) - e_{ESARIMA,t}^2(3) \end{aligned}$$

y

$$\begin{aligned} f &: \mathbb{R} \rightarrow \mathbb{R} \\ f(x) &= x \end{aligned}$$

para expresar H_{01}^3 en términos de H_0^G . Cuando, en cambio, consideramos

$$H_{02}^3 : PCM_{Consensus}^3 \geq PCM_{EEE}^3 \geq PCM_{ESARIMA}^3$$

basta tomar

$$\begin{aligned} X_{1t} &= \left(Y_{Consensus,t}^f(3) \right)^2 - \left(Y_{EEE,t}^f(3) \right)^2 \\ X_{2t} &= \left(Y_{Consensus,t}^f(3) \right)^2 - \left(Y_{ESARIMA,t}^f(3) \right)^2 \\ X_{3t} &= \left(Y_{EEE,t}^f(3) \right)^2 - \left(Y_{ESARIMA,t}^f(3) \right)^2 \end{aligned}$$

y

$$\begin{aligned} f &: \mathbb{R} \rightarrow \mathbb{R} \\ f(x) &= -x \end{aligned}$$

Un test para H_0^G es desarrollado por White (2000) bajo los supuestos que guían la teoría asintótica fuera de muestra desarrollada por West (1996). Si en vez de considerar esta teoría asintótica, usamos la propuesta por Giacomini y White (2006), podemos desarrollar un test similar al de White (2000). En efecto, tenemos que cuando el número de predicciones involucradas, P , tiende a infinito, se cumple que

$$\sqrt{P}(\bar{Z} - \mathbb{E}Z_t) \overset{A}{\rightsquigarrow} N(0, \Omega)$$

en que Ω representa a una matriz definida positiva que algunas veces es denominada como “matriz de varianzas-covarianzas de largo plazo”. Como resultado de la aplicación del teorema de la función continua para convergencia en distribución, tenemos que en la medida que el número de pronósticos P vaya a infinito, entonces

$$\underset{j \in \{1,2,\dots,l\}}{\text{Max}} \left[\sqrt{P} \frac{1}{P} \sum_{t=1}^P [Z_{jt} - \mathbb{E}Z_{jt}] \right]_{l \times 1} \xrightarrow{\mathcal{D}}_A \underset{k \in \{1,\dots,l\}}{\text{Max}} \{u_k\}$$

donde $\{u_1, u_2, \dots, u_l\}$ es un vector de dimensión l que sigue una distribución normal multivariada $N(0, \Omega)$.

A partir de este resultado es posible construir un test basado en el estadístico del máximo:

$$\underset{j \in \{1,2,\dots,l\}}{\text{Max}} \left[\sqrt{P} \frac{1}{P} \sum_{t=1}^P [Z_{jt} - \mathbb{E}Z_{jt}] \right]_{l \times 1}$$

No obstante, y como lo señalan Hansen (2005) y Romano y Wolf (2005), es intuitivamente preferible construir un test sobre un vector de variables Z_t estandarizado. Para ello procedemos de la siguiente manera. Consideremos la siguiente matriz diagonal \mathbf{D} , definida como

$$\mathbf{D}_{ii} = \sigma_i^{-1}; i = 1, \dots, l$$

en que

$$\sigma_i^2 = \Omega_{ii}; i = 1, \dots, l$$

entonces se cumple que

$$\sqrt{P}\mathbf{D}(\bar{Z} - \mathbb{E}Z_t) \overset{A}{\rightsquigarrow} N(0, \mathbf{D}\Omega\mathbf{D})$$

con el beneficio de que ahora

$$[\mathbf{D}\Omega\mathbf{D}]_{ii} = 1; \text{ para todo } i = 1, \dots, l$$

Desafortunadamente tenemos tres términos desconocidos: $\mathbb{E}Z_t$, \mathbf{D} y Ω . Notemos que la hipótesis nula H_0^G se puede expresar como

$$H_0^G : \mathbb{E}Z_t \leq 0$$

Esto muestra que hay una infinidad de vectores $\mathbb{E}Z_t$ que son coherentes con la hipótesis nula. En particular

$$\mathbb{E}Z_t = 0 \tag{2}$$

es un valor coherente y es el menos favorable a la alternativa, lo que asegura la construcción de un test con un tamaño acotado por el tamaño nominal. Por esta razón, y en forma coherente con la propuesta de White (2000) construimos nuestro test imponiendo (2).

También tenemos que \mathbf{D} y Ω son desconocidos, pero mediante el método de Newey y West (1987, 1994) por ejemplo, es posible construir un estimador consistente y definido positivo para Ω que denotaremos $\widehat{\Omega}$. Luego podemos definir un estimador consistente de \mathbf{D} como sigue

$$\widehat{\mathbf{D}}_{ii} = \left[\widehat{\Omega}_{ii} \right]^{-1/2}; i = 1, \dots, l$$

Luego se cumple que bajo la hipótesis nula

$$\sqrt{P}\widehat{\mathbf{D}}(\overline{Z}) \overset{A}{\rightsquigarrow} N(0, \widetilde{\Omega})$$

donde

$$\widetilde{\Omega} \equiv \mathbf{D}\Omega\mathbf{D}$$

Con esto, nuestro estadístico de interés se puede escribir como

$$\underset{j \in \{1, 2, \dots, l\}}{\text{Max}} \sqrt{P}\widehat{\mathbf{D}}\overline{Z}$$

en que $j \in \{1, 2, \dots, l\}$ representan a las distintas componentes del vector

$$\widehat{\mathbf{D}}\overline{Z}$$

Los valores críticos de nuestro estadístico pueden ser calculados a través de simulaciones de Monte Carlo, como lo sugiere White (2000), de acuerdo a los siguientes pasos

1. Calcular la siguiente descomposición de Cholesky sobre $\widehat{\mathbf{D}}\widehat{\Omega}\widehat{\mathbf{D}} = G'G$ con G triangular superior.
2. Definir un número de replicaciones N , que en el presente artículo corresponderá a $N = 1,000,000$.
3. Para cada una del millón de replicaciones se obtiene una realización independiente v de una normal multivariada $N(0, I_{l \times l})$.
4. Se define w como $w = G'v$, luego w es una realización independiente de $N(0, \widehat{\mathbf{D}}\widehat{\Omega}\widehat{\mathbf{D}})$.
5. Se define m como

$$m = \underset{i \in \{1, \dots, l\}}{\text{máx}} \{w_i\}$$

6. Luego se ordena el millón de términos m obtenidos y se definen los valores críticos como los correspondientes cuantiles.

Es importante recalcar que bajo la hipótesis nula genérica

$$\underset{j \in \{1, 2, \dots, l\}}{\text{Max}} \mathbb{E}Z_t \leq 0$$

mientras que bajo la hipótesis alternativa genérica se tendrá que

$$\underset{j \in \{1, 2, \dots, l\}}{\text{Max}} \mathbb{E}Z_t > 0$$

lo que implica que nos concentraremos en un test de una sola cola, cuya región de rechazo estará definida por una relación del tipo

$$\underset{j \in \{1, 2, \dots, l\}}{\text{Max}} \sqrt{\widehat{PDZ}} > c$$

Naturalmente también es posible construir de manera análoga un test para el estadístico del mínimo, lo cual es muy útil para el caso en el que la hipótesis nula se exprese como

$$H_0 : \mathbb{E}Z_t \leq 0$$

En las próximas líneas veremos que nos dicen estos test conjuntos basados en los estadísticos del máximo y del mínimo, cuando los aplicamos a las hipótesis de nuestro interés. Para todos los cálculos se utiliza la muestra desde Dic-2003 hasta Dic-2009 (73 datos).

Errores Cuadráticos: Fuente Variable y Horizonte Fijo Bajo el criterio del ECM se desea probar el cumplimiento de los siguientes perfiles :

$$H_{01}^1 : ECM_{Bloomberg}^1 \leq ECM_{Consensus}^1 \leq ECM_{EEE}^1 \leq ECM_{ESARIMA}^1$$

$$H_{01}^3 : ECM_{Consensus}^3 \leq ECM_{EEE}^3 \leq ECM_{ESARIMA}^3$$

lo que implica el cumplimiento simultáneo de seis desigualdades para el caso en que $h = 1$ y de tres desigualdades para el caso $h = 3$. Los resultados se presentan en la Tabla 11, desde donde se aprecia que no es posible rechazar la hipótesis nula, ya que los estadísticos calculados se encuentran en el extremo izquierdo de la distribución, corroborando los resultados obtenidos usando las reglas 1 y 2.

Tabla 11: Percentiles fdp (*).

h=1				h=3			
Percentil		Percentil		Percentil		Percentil	
Calculado	-1.953	50 %	0.695	Calculado	-0.817	50 %	0.603
1 %	-1.184	60 %	0.909	1 %	-1.141	60 %	0.817
5 %	-0.646	70 %	1.141	5 %	-0.652	70 %	1.049
10 %	-0.355	80 %	1.415	10 %	-0.388	80 %	1.326
20 %	0.000	90 %	1.803	20 %	-0.063	90 %	1.716
30 %	0.259	95 %	2.122	30 %	0.177	95 %	2.045
40 %	0.483	99 %	2.735	40 %	0.394	99 %	2.661

(*) Fuente: Elaboración propia.

Proyecciones Cuadráticas: Fuente Variable y Horizonte Fijo Bajo el criterio del PCM se desea probar el cumplimiento de los siguientes perfiles:

$$H_{02}^1 : PCM_{Bloomberg}^1 \geq PCM_{Consensus}^1 \geq PCM_{EEE}^1 \geq PCM_{ESARIMA}^1$$

$$H_{02}^3 : PCM_{Consensus}^3 \geq PCM_{EEE}^3 \geq PCM_{ESARIMA}^3$$

Para el caso en que $h = 1$ entonces deben satisfacerse seis condiciones para el cumplimiento conjunto de la desigualdad anterior. Para el caso en que $h = 3$, la dimensión del problema se reduce a tres desigualdades.

La forma de estimar la fdp es similar a la del caso anterior, con la única diferencia que el estadístico se construye sobre el mínimo de las componentes y ya no sobre el máximo. Con esto, bajo la nula se espera que el mínimo de las diferencias de PCM sea positivo. Bajo la alternativa, en cambio, el mínimo debiese ser negativo, lo que implica que nos concentraremos en un test de una sola cola, cuya región de rechazo estará definida por una relación del tipo

$$\underset{j \in \{1, 2, \dots, l\}}{\text{Min}} \sqrt{P\hat{D}\bar{Z}} < c$$

con \bar{Z} representando al vector de las diferencias en los PCM estimados.

Los resultados de la aplicación de este test se encuentran en la Tabla 12. Estos indican que al 5% de significancia se rechaza la hipótesis nula de optimalidad desde la perspectiva de las PCM para los dos horizontes analizados, ya que los estadísticos calculados son menores que el cuantil correspondiente al 5% de probabilidad acumulada sobre la distribución bajo la hipótesis nula, confirmando lo encontrado en las secciones anteriores con las reglas 1 y 2. En consecuencia, el supuesto de optimalidad bajo pérdida cuadrática es rechazado para el perfil de proyecciones.

Tabla 12: Percentiles fdp (*).

h=1				h=3			
Percentil		Percentil		Percentil		Percentil	
Calculado	-2.337	50 %	-0.436	Calculado	-2.255	50 %	-0.716
1 %	-2.608	60 %	-0.206	1 %	-2.784	60 %	-0.504
5 %	-1.963	70 %	0.033	5 %	-2.162	70 %	-0.281
10 %	-1.621	80 %	0.313	10 %	-1.835	80 %	-0.025
20 %	-1.211	90 %	0.703	20 %	-1.446	90 %	-0.323
30 %	-0.915	95 %	1.02	30 %	-1.167	95 %	0.609
40 %	-0.665	99 %	1.611	40 %	-0.932	99 %	1.142

(*) Fuente: Elaboración propia.

Errores Cuadráticos: Fuente Fija y Horizonte Variable Bajo el criterio del ECM se desea probar el cumplimiento de los siguientes perfiles :

$$H_{03}^{Consensus} : ECM_{Consensus}^{h=1} \leq ECM_{Consensus}^{h=2} \leq \dots \leq ECM_{Consensus}^{h=6}$$

$$H_{03}^{EEE} : ECM_{EEE}^{h=1} \leq ECM_{EEE}^{h=3} \leq ECM_{EEE}^{h=12}$$

$$H_{03}^{ESARIMA} : ECM_{ESARIMA}^{h=1} \leq ECM_{ESARIMA}^{h=2} \leq \dots \leq ECM_{ESARIMA}^{h=6} \leq ECM_{ESARIMA}^{h=12}$$

es decir, se desea evaluar si cada fuente de información entrega una estructura temporal de pronósticos a distintos horizontes compatible con optimalidad bajo pérdida cuadrática.

Los distintos percentiles correspondientes a la distribución del estadístico bajo la hipótesis nula, así como el cálculo del estadístico propiamente tal, se presentan en las siguientes tablas 13,14 y 15. Los estadísticos son todos negativos y se encuentran muy alejados hacia la izquierda en la distribución, indicando que no es posible rechazar la hipótesis nula de optimalidad del perfil de ECM.

Tabla 13: Percentiles *fdp* test Errores Cuadráticos Consensus Forecasts. (*).

Percentil		Percentil	
Calculado	-1.453	50 %	0.842
1 %	-0.671	60 %	1.035
5 %	-0.268	70 %	1.246
10 %	-0.042	80 %	1.503
20 %	0.245	90 %	1.869
30 %	0.463	95 %	2.181
40 %	0.655	99 %	2.786

(*) Fuente: Elaboración propia.

Tabla 14: Percentiles *fdp* test Errores Cuadráticos EEE (*).

Percentil		Percentil	
Calculado	-3.001	50 %	0.482
1 %	-1.478	60 %	0.702
5 %	-0.914	70 %	0.940
10 %	-0.611	80 %	1.222
20 %	-0.239	90 %	1.618
30 %	0.031	95 %	1.946
40 %	0.264	99 %	2.574

(*) Fuente: Elaboración propia.

Tabla 15: Percentiles *fdp* test Errores Cuadráticos ESARIMA (*).

Percentil		Percentil	
Calculado	-1.862	50 %	0.802
1 %	-1.066	60 %	1.015
5 %	-0.529	70 %	1.243
10 %	-0.242	80 %	1.514
20 %	0.111	90 %	1.896
30 %	0.368	95 %	2.212
40 %	0.592	99 %	2.818

(*) Fuente: Elaboración propia.

Pronósticos Cuadráticos: Fuente Fija y Horizonte Variable Bajo el criterio del PCM se desea probar el cumplimiento de las siguientes hipótesis conjuntas:

$$H_{04}^{Consensus} : PCM_{Consensus}^{h=1} \geq PCM_{Consensus}^{h=2} \geq \dots \geq PCM_{Consensus}^{h=6}$$

$$H_{04}^{EEE} : PCM_{EEE}^{h=1} \geq PCM_{EEE}^{h=3} \geq PCM_{EEE}^{h=12}$$

$$H_{04}^{ESARIMA} : PCM_{ESARIMA}^{h=1} \geq PCM_{ESARIMA}^{h=2} \geq \dots \geq PCM_{ESARIMA}^{h=6} \geq PCM_{ESARIMA}^{h=12}$$

en que $h = 1, \dots, 12$ denotan los horizontes de proyección.

Es decir, se desea evaluar si cada fuente de información entrega una estructura temporal de pronósticos a distintos horizontes compatible con optimalidad bajo pérdida cuadrática, pero en términos de un ordenamiento sobre pronósticos cuadráticos medio. Observamos tanto en la Tabla 16 como en la 17 y 18 que el estadístico calculado cae completamente fuera de la región de rechazo para las tres fuentes de pronósticos consideradas. Para Consensus el estadístico calculado entrega un valor de 1.505, mientras que la región de rechazo está dada por valores menores que -1.770 al 10 % de significancia. Para la EEE y las proyecciones

ESARIMA, los estadísticos calculados son 0.902 y -0.129 respectivamente, mientras que las zonas de rechazo están dadas por valores inferiores a -1.632 y -1.922 respectivamente.

Tabla 16: Percentiles *fdp* test Proyecciones Cuadráticas Consensus Forecasts (*).

Percentil		Percentil	
Calculado	1.505	50 %	-0.605
1 %	-2.741	60 %	-0.377
5 %	-2.104	70 %	-0.135
10 %	-1.770	80 %	0.146
20 %	-1.366	90 %	0.537
30 %	-1.078	95 %	0.857
40 %	-0.832	99 %	1.443

(*) Fuente: Elaboración propia.

Tabla 17: Percentiles *fdp* test Proyecciones Cuadráticas EEE (*).

Percentil		Percentil	
Calculado	0.902	50 %	-0.543
1 %	-2.568	60 %	-0.336
5 %	-1.953	70 %	-0.116
10 %	-1.632	80 %	0.136
20 %	-1.251	90 %	0.484
30 %	-0.981	95 %	0.767
40 %	-0.753	99 %	1.286

(*) Fuente: Elaboración propia.

Tabla 18: Percentiles *fdp* test Proyecciones Cuadráticas ESARIMA (*).

Percentil		Percentil	
Calculado	-0.129	50 %	-0.868
1 %	-2.838	60 %	-0.668
5 %	-2.235	70 %	-0.456
10 %	-1.922	80 %	-0.211
20 %	-1.552	90 %	0.119
30 %	-1.291	95 %	0.389
40 %	-1.070	99 %	0.884

(*) Fuente: Elaboración propia.

Estos resultados confirman los obtenidos en secciones anteriores con la aplicación de las reglas 1 y 2 y no pueden rechazar una estructura temporal, de cada una de las fuentes de proyección, que es coherente con optimalidad bajo pérdida cuadrática.

El lector podría preguntarse si era necesario confirmar los resultados obtenidos con las reglas 1 y 2 usando un test conjunto. Después de todo, estas tres estrategias han entregado resultados similares. En la siguiente sección mostraremos con simulaciones por qué es importante la construcción de un test conjunto.

6. Análisis de los Resultados

6.1. Necesidad de un Test Conjunto

Cuando el interés del investigador se centra en una hipótesis conjunta que se compone de múltiples hipótesis univariadas se pueden adoptar básicamente dos estrategias para realizar inferencia estadística. La primera consiste en construir múltiples tests para cada una de las hipótesis univariadas y luego, mediante alguna regla, obtener una conclusión de aceptación o rechazo de la hipótesis conjunta bajo evaluación. Las reglas 1 y 2 presentadas en secciones anteriores son ejemplos de estos procedimientos. No obstante la simpleza que ofrecen estas estrategias, hay dos críticas fundamentales que se relacionan con el adecuado control del error de tipo I de la prueba conjunta. En efecto, tanto la regla 1 como la 2 (cotas de Bonferroni) utilizadas en este artículo pueden presentar serias dificultades en el control del error de tipo I. Para analizar un caso en particular en el que esto puede ocurrir para la regla 1, pensemos en la siguiente hipótesis nula multivariada:

$$\begin{aligned} H_0 & : \vec{\mu} = 0_{N \times N} \\ \vec{\mu} & = \mathbb{E}X_t \end{aligned}$$

En otras palabras deseamos testear que las N componentes de la esperanza del vector aleatorio X_t sean cero. La hipótesis alternativa plantea que al menos una componente de $\vec{\mu}$ es mayor que cero:

$$H_A : \exists i \in \{1, \dots, N\} \text{ tal que } \mu_i > 0$$

Supongamos que disponemos de una colección de tests

$$T_i, i = 1, \dots, N$$

que son una función de la muestra y que están diseñados para testear la siguiente hipótesis nula

$$H_0^{(i)} : \mu_i = 0$$

con respecto a la alternativa

$$H_A^{(i)} : \mu_i > 0$$

de manera que cada test T_i rechaza la nula $H_0^{(i)}$ a un determinado nivel α cuando, evaluados en la muestra disponible de datos D ,

$$T_i(D) > c$$

Evaluemos el tamaño del error de tipo I de la regla 1 bajo el supuesto de independencia del vector

$$\vec{T} = (T_1, T_2, \dots, T_N)'$$

En este caso tenemos que

$$\Pr(\exists i \in \{1, \dots, N\} \text{ tal que } T_i(D) > c | H_0) = \Pr\left(\sum_{i=1}^N W_i > 0 | H_0\right)$$

en que

$$W_i = \left\{ \begin{array}{ll} 1 & \text{si } T_i(D) > c \\ 0 & \text{si no} \end{array} \right\}$$

Luego W_i es una variable aleatoria Bernoulli(p) de parámetro

$$p \leq \alpha$$

luego, bajo la nula $\sum_{i=1}^N W_i$ es una binomial de parámetros N, p

$$\sum_{i=1}^N W_i \rightsquigarrow B(N, p)$$

De esta forma:

$$\begin{aligned} \Pr\left(\sum_{i=1}^N W_i > 0 | H_0\right) &= 1 - \Pr\left(\sum_{i=1}^N W_i = 0 | H_0\right) \\ &= 1 - \Pr(T_i(D) \leq c \forall i \in \{1, \dots, N\} | H_0) \\ &= 1 - (1 - p)^N \rightarrow 1 \text{ cuando } N \rightarrow \infty \end{aligned}$$

Es decir la regla 1 pierde completamente el control del error de tipo I en la medida que la dimensión del problema tiende a infinito y bajo el supuesto de independencia del vector de tests individuales.

En el caso de la regla 2, esto es del uso de las cotas de Bonferroni, su problema usual es plantear un test a veces demasiado conservador puesto que se cumple la siguiente desigualdad

$$\Pr(\exists i \in \{1, \dots, N\} \text{ tal que } T_i(D) > \tilde{c} | H_0) \leq \sum_{i=1}^N \Pr(T_i(D) > \tilde{c} | H_0)$$

en el caso de la regla 2 se escoge \tilde{c} tal que

$$\Pr(T_i(D) > \tilde{c} | H_0) = \frac{\alpha}{N} \forall i \in \{1, \dots, N\}$$

luego, estas cotas garantizan que

$$\Pr(\exists i \in \{1, \dots, N\} \text{ tal que } T_i(D) > \tilde{c} | H_0) \leq \sum_{i=1}^N \frac{\alpha}{N} = \alpha$$

pero nada asegura que esta relación no se cumpla con desigualdad estricta y que en realidad el tamaño del test sea bastante inferior al tamaño deseado.

La tabla 19 a continuación, muestra el error de tipo I cometido por el test del máximo y las reglas 1 y 2 cuando el problema es testear

$$\begin{aligned} H_0 &: \vec{\mu} = 0_{2 \times 2} \\ \vec{\mu} &= \mathbb{E}X_t \end{aligned}$$

versus

$$H_A : \exists i \in \{1, 2\} \text{ tal que } \mu_i > 0$$

Los valores en cada tabla corresponden a los tamaños empíricos calculados en 100,000 simulaciones de Monte Carlo en el que los parámetros verdaderos están fijos en cero y corresponden a la esperanza de una normal bivariada con varianza V de diagonal unitaria y en el que se prueban varias configuraciones de covarianzas. El tamaño nominal de la prueba conjunta se fija en 0.10, por lo que se espera que el tamaño empírico se aproxime tanto como sea posible a 0.10. Los correspondientes tests individuales son test t tradicionales.

Tabla 19: Error de tipo I con tamaño nominal del 10% (*)

Evaluación del Test del Máximo y Reglas 1 y 2 Cuando N=2

Correlación	Test del Máximo	Regla 1	Regla 2
r=-0.90	0.100	0.200	0.099
r=-0.50	0.100	0.201	0.098
r=0.00	0.100	0.188	0.098
r=0.50	0.100	0.168	0.091
r=0.90	0.100	0.133	0.069

(*) Fuente: Elaboración propia.

Se observa que el test del máximo obtiene un error de tipo I perfecto, mientras que las reglas de tipo 1 y 2 tienen errores de tipo I excesivamente altos y bajos respectivamente, lo que justifica la construcción de la prueba conjunta propuesta.

6.2. Fuente de Ineficiencias

La ineficiencia encontrada en este artículo plantea que si tenemos dos pronósticos distintos e insesgados $y_{t-h}^{(1)}(h)$ y $y_{t-h}^{(2)}(h)$ entonces

$$\begin{aligned} Y_t &= y_{t-h}^{(1)}(h) + e_{t-h}^{(1)}(h) \\ Y_t &= y_{t-h}^{(2)}(h) + e_{t-h}^{(2)}(h) \end{aligned}$$

luego

$$\begin{aligned} V(Y_t) &= V(y_{t-h}^{(1)}(h)) + V(e_{t-h}^{(1)}(h)) + 2Cov(y_{t-h}^{(1)}(h), e_{t-h}^{(1)}(h)) \\ V(Y_t) &= V(y_{t-h}^{(2)}(h)) + V(e_{t-h}^{(2)}(h)) + 2Cov(y_{t-h}^{(2)}(h), e_{t-h}^{(2)}(h)) \end{aligned}$$

o equivalentemente

$$\begin{aligned} \mathbb{E}(Y_t^2) &= \mathbb{E}(y_{t-h}^{(1)}(h))^2 + \mathbb{E}(e_{t-h}^{(1)}(h))^2 + 2\mathbb{E}(y_{t-h}^{(1)}(h)e_{t-h}^{(1)}(h)) \\ \mathbb{E}(Y_t^2) &= \mathbb{E}(y_{t-h}^{(2)}(h))^2 + \mathbb{E}(e_{t-h}^{(2)}(h))^2 + 2\mathbb{E}(y_{t-h}^{(2)}(h)e_{t-h}^{(2)}(h)) \end{aligned}$$

Si

$$\mathbb{E}(e_{t-h}^{(1)}(h))^2 < \mathbb{E}(e_{t-h}^{(2)}(h))^2$$

pero también

$$\mathbb{E}(y_{t-h}^{(1)}(h))^2 \leq \mathbb{E}(y_{t-h}^{(2)}(h))^2$$

entonces debe tenerse que

$$\mathbb{E}(y_{t-h}^{(1)}(h))^2 + \mathbb{E}(e_{t-h}^{(1)}(h))^2 < \mathbb{E}(y_{t-h}^{(2)}(h))^2 + \mathbb{E}(e_{t-h}^{(2)}(h))^2$$

o equivalentemente

$$\mathbb{E}(Y_t^2) - 2\mathbb{E}(y_{t-h}^{(1)}(h)e_{t-h}^{(1)}(h)) < \mathbb{E}(Y_t^2) - 2\mathbb{E}(y_{t-h}^{(2)}(h)e_{t-h}^{(2)}(h))$$

que es lo mismo que

$$\mathbb{E}(y_{t-h}^{(1)}(h)e_{t-h}^{(1)}(h)) > \mathbb{E}(y_{t-h}^{(2)}(h)e_{t-h}^{(2)}(h))$$

o bien

$$\mathbb{E}(y_{t-h}^{(1)}(h)e_{t-h}^{(1)}(h)) - \mathbb{E}(y_{t-h}^{(2)}(h)e_{t-h}^{(2)}(h)) > 0$$

lo que implica que al menos uno de los términos de la izquierda en la última expresión es no nulo, lo que es claramente una violación a una condición de ortogonalidad implicada por la optimalidad bajo pérdida cuadrática⁴.

La exploración de las tablas 9 y 10 nos indican que una situación como la descrita anteriormente, es decir

$$\mathbb{E}(e_{t-h}^{(1)}(h))^2 < \mathbb{E}(e_{t-h}^{(2)}(h))^2$$

pero también

$$\mathbb{E}(y_{t-h}^{(1)}(h))^2 \leq \mathbb{E}(y_{t-h}^{(2)}(h))^2$$

sucede para los siguientes casos:

1. Comparaciones 1 mes hacia adelante entre Bloomberg y ESARIMA, Consensus y ESARIMA y EEE y ESARIMA.

⁴Notemos que los horizontes predictivos no tienen por que ser los mismos para esta propiedad.

2. Comparaciones 3 meses hacia adelante entre EEE y ESARIMA.
3. Comparaciones 12 meses hacia adelante entre EEE y ESARIMA.

Las siguientes tablas 20 y 21 nos muestran los resultados de estas 6 comparaciones y de los 8 términos individuales que forman parte de estas 6 comparaciones

Tabla 20: Resultados test de eficiencia débil. Muestra restringida

Pruebas Diferenciales			Muestra: Dic-2003 – Dic-2009 (73 datos)			
EEE y ESARIMA	CE y ESARIMA	Bloomberg y ESARIMA	ESARIMA	EEE	C. Economics	Bloomberg
h=1 0.468251 ^{†††} (0.197368)	0.477899 ^{††} (0.236829)	0.429758 [†] (0.287424)	-0.215707 (0.401743)	0.252543 (0.325289)	0.262192 (0.313666)	0.214051 (0.217124)
h=3 1.615738 ^{††} (0.682161)			-1.038374 (1.531720)	0.577364 (1.240103)		- -
h=12 11.96060 [†] (6.362926)		- -	-8.719918 (7.222034)	3.240685 [†] (1.906931)		- -

(*) Entre paréntesis se presenta la desviación estándar. (†) Significativo al 10 %.
(††) Significativo al 5 % (†††) Significativo al 1 % Fuente: Elaboración propia.

Tabla 21: Resultados Test de Eficiencia Débil. Muestra completa

Pruebas Diferenciales			Muestra completa			
EEE y ESARIMA	CE y ESARIMA	Bloomberg y ESARIMA	ESARIMA	EEE	C. Economics.	Bloomberg
h=1 0.359616 ^{†††} (0.147259)	0.377310 ^{††} (0.176984)	0.278649 ^{††} (0.154895)	-0.237246 (0.289309)	0.122370 (0.236084)	0.130943 (0.233661)	0.161712 (0.149990)
h=3 1.230501 ^{††} (0.529064)			-0.982160 (1.108606)	0.202078 (0.944347)		- -
h=12 8.929338 ^{††} (5.095317)		- -	-6.239123 (5.348414)	2.160792 [†] (1.589705)		- -

(*) Entre paréntesis se presenta la desviación estándar. (†) Significativo al 10 %.
(††) Significativo al 5 % (†††) Significativo al 1 % Fuente: Elaboración propia.

En estas tablas puede apreciarse que efectivamente la hipótesis

$$H_o : \mathbb{E}(y_{t-h}^{(1)}(h)e_{t-h}^{(1)}(h)) - \mathbb{E}(y_{t-h}^{(2)}(h)e_{t-h}^{(2)}(h)) = 0$$

es rechazada en todos los casos anteriormente destacados en los puntos 1-3 en favor de la alternativa

$$H_A : \mathbb{E}(y_{t-h}^{(1)}(h)e_{t-h}^{(1)}(h)) - \mathbb{E}(y_{t-h}^{(2)}(h)e_{t-h}^{(2)}(h)) > 0$$

De esta forma los paneles del lado izquierdo en las tablas 20-21 indican que hay información no eficientemente utilizada por alguno de los pronósticos considerados en la evaluación. Interesantemente, los paneles del lado derecho de las mismas tablas solo dan cuenta de una ineficiencia estadísticamente significativa para la EEE en proyecciones a 12 meses. No obstante, también se debe apreciar que al usar niveles de significancia algo más holgados, los mismos paneles del lado derecho de la tabla 21 muestran un rechazo de la hipótesis nula de eficiencia para las proyecciones ESARIMA a 12 meses a un nivel de significancia del 13 % en tests de una sola cola. Similarmente, la tabla 21 también muestra un rechazo de la eficiencia de las proyecciones de Bloomberg a una sola cola al 15 %. Adicionalmente, y en un resultado no mostrado simplemente por abreviar espacio, es posible mostrar que la eficiencia de las proyecciones de Consensus a horizontes de 3,4,5 y 6 meses es rechazada en la muestra completa con un nivel de significancia del 5 %. Esto configura un escenario en el que surge evidencia de ineficiencia que es transversal al conjunto de proyecciones que se considere y al horizonte predictivo. Todo lo cual sugiere que existe espacio para modificar los conjuntos de proyecciones de manera de hacerlos más precisos aún. Estos resultados son coherentes con los mostrados por Bentancor y Pincheira (2010) y Pincheira y Fernández (2010) para las proyecciones de la EEE en Chile y las de Consensus en varios países.

7. Resumen y conclusiones

En este trabajo hemos analizado la optimalidad bajo pérdida cuadrática de una serie de pronósticos de inflación para Chile. Nuestro análisis ha explotado propiedades de optimalidad que plantean restricciones de desigualdad sobre la estructura temporal de las proyecciones, lo cual, de acuerdo a nuestra revisión de la literatura, no se ha hecho anteriormente para el caso chileno. Si bien algunas de las restricciones estudiadas corresponden a propiedades vastamente conocidas que relacionan a los Errores Cuadrático Medio (ECM) de proyecciones construidas en distintos momentos o para distintos horizontes, otras propiedades evaluadas se refieren a la estructura temporal de los pronósticos propiamente tales y no de sus errores. Estas propiedades han sido recientemente derivadas por Patton y Timmerman (2010).

Una segunda contribución de nuestro trabajo se refiere a la manera en la cual estas propiedades sobre la estructura temporal de las proyecciones es evaluada. Patton y Timmermann (2010) diseñan test basados en las contribuciones de Wolak (1987,1989) que requieren la solución de uno o varios problemas de programación cuadrática convexa. En cambio en este trabajo nosotros diseñamos tests basados en los estadísticos del mínimo y del máximo que no requieren solucionar un problema de optimización. Estos test se inspiran en los resultados de White (2000), Pincheira (2006), Hansen (2005) y Romano y Wolf (2005).

Es importante destacar la utilidad que puede tener un análisis de eficiencia predictiva. En efecto, independientemente del conjunto de información y función de pérdida que posea el pronosticador que generó las proyecciones bajo análisis, es perfectamente razonable preguntarse si esas proyecciones pueden mejorarse bajo el criterio que un economista desee. Así, una evaluación de optimalidad, puede servir de orientación para entender de que manera una proyección pueda ser inmediatamente ajustada y mejorada de acuerdo a la función de pérdida de cada economista.

Desde el punto de vista empírico, en este trabajo aplicamos un test conjunto de superioridad predictiva a las proyecciones de inflación chilena disponibles a través de las siguientes fuentes: Bloomberg, Consensus Economics, la Encuesta de Expectativas Económicas del Banco Central de Chile y un promedio de modelos ESARIMA. Los resultados indican que en términos de Error Cuadrático Medio no es posible rechazar la hipótesis nula de optimalidad para el perfil conjunto de proyecciones. No obstante, cuando el test es aplicado a las Proyecciones al Cuadrado, en línea con la propuesta de Patton y Timmermann (2010), el test conjunto rechaza la hipótesis de optimalidad. El análisis de estos resultados sugiere que existe una violación a una condición de ortogonalidad implicada por la optimalidad bajo pérdida cuadrática, y que esta violación es transversal al conjunto de proyecciones bajo consideración y al horizonte predictivo, todo lo cual sugiere la existencia de espacio para perfeccionar el perfil del proyecciones.

8. Referencias

1. Andersson, M., Karlsson, G. and J. Svensson (2007): The Riksbank's Forecasting Performance. *Economic Review* **3**: 59-75.
2. Ang, A., G. Bekaert y M. Wei (2007). Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?. *Journal of Monetary Economics* **54**(4): 1163-1212.
3. Bentancor, A. y P. Pincheira (2010). Predicción de Errores de Proyección de Inflación en Chile. *El Trimestre Económico* **LXXVII**(305): 129-154.
4. Bowles, C., R. Friz, V. Genre, G. Kenny, A. Meyler y T. Rautanen (2007). The ECB Survey of Professional Forecasters (SPF) - A Review After Eight Year's Experience. Occasional Paper Series 59, European Central Bank.
5. Capistrán, C. y G. López-Moctezuma (2010). Las Expectativas Macroeconómicas de los Especialistas. Una Evaluación de Pronósticos de Corto Plazo en México. *El Trimestre Económico* **LXXVII**(306): 275-312.

6. Chumacero, R. (2001). Empirical Analysis of Systematic Errors in Chilean GDP Forecasts. *Journal of Forecasting* **20**(1): 37-45.
7. Croushore, D. (2010). An Evaluation of Inflation Forecasts from Surveys Using Real-Time Data. *The B.E. Journal of Macroeconomics* **10**(1).
8. Diebold, F.X. y R. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* **13**(3): 253-63.
9. Elliot, G. y A. Timmermann (2008). Economic Forecasting. *Journal of Economic Literature* **46**(1): 3-56.
10. Giacomini, R. y H. White (2006). Tests of Conditional Predictive Ability. *Econometrica* **74**(6): 1545-1578.
11. Groen J, Kapetanios G. y S. Price (2009) A real time evaluation of Bank of England forecasts of inflation and growth. *International Journal of Forecasting* **25**: 74-80
12. Hansen, P.R. (2005). A Test of Superior Predictive Ability. *Journal of Business & Economic Statistics* **23**: 365-380.
13. Joutz, F. y H.O. Stekler (2000). An Evaluation of the Prediction of the Federal Reserve. *International Journal of Forecasting* **16**: 17 – 38.
14. Loungani, P. (2001). How Accurate are Private Sector Forecasts? Cross-country Evidence from Consensus Forecast of Output Growth. *International Journal of Forecasting* **17**(3): 419-432.
15. Newey, W. y K. West (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* **55**(3): 703-08.
16. Newey, W. y K. West (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* **61**(4): 631-53.
17. Patton, A. y A. Timmermann (2010). New Tests of Forecast Optimality Across Multiple Horizons. Working Paper, Univeristy of California, San Diego.
18. Pedersen, M. (2010). "Una Nota Introductoria a la Encuesta de Expectativas Económicas..^Estudio Económico Estadístico 82, Banco Central de Chile.
19. Pincheira, P. (2006). Shrinkage Based Tests of the Martingale Difference Hypothesis. Working Paper 376, Central Bank of Chile.
20. Pincheira, P. (2010). A Real Time Evaluation of the Central Bank of Chile GDP Growth Forecasts. *Money Affairs* Vol XXIII N°1, 37-73.
21. Pincheira, P. y R. Alvarez (2009). Evaluation of Short-Run Inflation Forecasts and Forecasters in Chile. *Money Affairs* Vol XXII N°2:159-180.
22. Pincheira, P. y N. Fernández (2010). Jaque Mate a las Proyecciones de Inflación de Consenso. *Manuscrito Banco Central de Chile*.
23. Pincheira, P. y A. García (2009). Forecasting Inflation in Chile with an Accurate Benchmark. Working Paper 514, Central Bank of Chile.
24. Romano, J.P. y M. Wolf (2005). Stepwise Multiple Testing as Formalized Data Sooping. *Econometrica* **73**(4): 1237-1282.
25. West, K. (1996). Asymptotic Inference About Predictive Ability. *Econometrica* **64**(5): 1067-84.

26. White, H. (2000). A Reality Check for Data Snooping. *Econometrica* **68**(5): 10987-1126.
27. Wolak, F.A. (1987). An Exact Test for Multiple Inequality and Equality Constraints in the Linear Regression Model. *Journal of the American Statistical Association* **82**: 782-793.
28. Wolak, F.A. (1989). Testing Inequality Constraints in Linear Econometric Models. *Journal of Econometrics* **31**: 205-235.

**Documentos de Trabajo
Banco Central de Chile**

**Working Papers
Central Bank of Chile**

NÚMEROS ANTERIORES

PAST ISSUES

La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: www.bcentral.cl/esp/estpub/estudios/dtbc. Existe la posibilidad de solicitar una copia impresa con un costo de \$500 si es dentro de Chile y US\$12 si es para fuera de Chile. Las solicitudes se pueden hacer por fax: (56-2) 6702231 o a través de correo electrónico: bcch@bcentral.cl.

Working Papers in PDF format can be downloaded free of charge from: www.bcentral.cl/eng/stdpub/studies/workingpaper. Printed versions can be ordered individually for US\$12 per copy (for orders inside Chile the charge is Ch\$500.) Orders can be placed by fax: (56-2) 6702231 or e-mail: bcch@bcentral.cl.

DTBC – 619 Marzo de 2011
**The Optimal Inflation Tax in the Presence of Imperfect Deposit –
Currency Substitution**
Eduardo Olaberría

DTBC – 618 Marzo de 2011
**El Índice de Cartera Vencida como Medida de Riesgo de Crédito:
Análisis y Aplicación al caso de Chile**
Andrés Sagner

DTBC – 617 Marzo 2011
Estimación del Premio por Riesgo en Chile
Francisca Lira y Claudia Sotz

DTBC – 616 Marzo 2011
**Uso de la aproximación TIR/Duración en la estructura de tasas:
resultados cuantitativos bajo Nelson – Siegel**
Rodrigo Alfaro y Juan Sebastián Becerra

DTBC – 615 Marzo de 2011
**Chilean Export Performance: the Rol of Intensive and Extensive
Margins**
Matías Berthelom

DTBC – 614 Febrero 2011
**Does Lineariry of the Dynamics of Inflation Gap and
Unemployment Rate Matter?**
Roque Montero

DTBC – 613	Febrero 2011
Modeling Copper Price: Regime Switching Approach Javier García – Cicco y Roque Montero	
DTBC – 612	Febrero 2011
Riding the Roller Coaster: Fiscal Policies of Non-renewable Resources Exporters in Latin America and the Caribbean. Mauricio Villafuerte Pablo López-Murphy y Rolando Ossowski	
DTBC – 611	Febrero 2011
Seigniorage and Distortionary Taxation in a Model with Heterogeneous Agents and Idiosyncratic Uncertainty Sofía Bauducco	
DTBC – 610	Febrero 2011
Stress Tests for Banking Sector: A technical Note Rodrigo Alfaro y Andrés Sagner	
DTBC – 609	Febrero 2011
Riding the Roller Coaster: Fiscal Policies of Nonrenewable Resources Exporters in Latin American and the Caribbean Mauricio Villafuerte, Pablo López – Murphy and Rolando Ossowski	
DTBC – 608	Febrero 2011
Floats, Pegs and the Transmission of Fiscal Policy Giancarlo Corsetti, Keith Kuester and Gernot J. Müller	
DTBC – 607	Enero 2011
A Bunch of Models, a Bunch of Nulls and Inference About Predictive Ability Pablo Pincheira	
DTBC – 606	Enero 2011
College Risk and Return Gonzalo Castex	
DTBC – 605	Enero 2011
Determinants of Export Diversification Around The World: 1962 – 2000 Manuel R. Agosin, Roberto Álvarez y Claudio Bravo-Ortega	