

Nowcasting del PIB de Uruguay

Evaluación del desempeño de los modelos en las crisis de
2002 y 2020

Maximiliano Mateauda Espinosa

Programa de Maestría en Economía de la Facultad de Ciencias
Económicas, Universidad de la República

Directora de tesis: Bibiana Lanzilotta

Codirector de tesis: Guillermo Carlomagno

Montevideo - Uruguay

Enero de 2022

INTEGRANTES DEL TRIBUNAL DE DEFENSA DE TESIS

Montevideo - Uruguay

Enero de 2022

AGRADECIMIENTOS

A mis tutores Bibiana y Guillermo, por guiarme en este proceso, por la dedicación y las valiosas contribuciones académicas.

A mis compañeras y compañeros de la Maestría y el trabajo, por los intercambios y compañía durante este tiempo.

A mi familia, por el apoyo incondicional.

A Rocío, por los valiosos aportes, su compañía y paciencia.

RESUMEN

Conocer el estado de la situación económica en tiempo real es un insumo fundamental para la toma de decisiones de los agentes económicos, principalmente los hacedores de política. Debido a que en Uruguay las estadísticas de Cuentas Nacionales tienen un rezago aproximado de 80 días, resulta necesario desarrollar alternativas para obtener información oportuna. Esto se vuelve aún más relevante en momentos donde la actividad económica sufre cambios abruptos. En este marco, el primer objetivo del trabajo consistió en estimar modelos de Nowcasting para predecir el PIB trimestral uruguayo mediante indicadores mensuales. Se aplicaron las metodologías de Modelos de Factores Dinámicos (FD) y modelos basados en Ecuaciones Puentes (EP). En segundo lugar, se analizó la contribución de cada indicador mensual a las revisiones de los pronósticos. Por último, se evaluó su capacidad predictiva en la crisis del 2002 y la de 2020.

Se concluye que los modelos propuestos tienen un mejor desempeño respecto a los benchmark usados como referencia, y que, con excepción de uno, las predicciones mejoran conforme incorporan la nueva información mensual. Al comparar entre metodologías, los EP son los que presentan mejor capacidad predictiva. Se encuentra que hay cinco variables que explican más del 50% de las revisiones absolutas de los pronósticos: El Índice de Volumen Físico de la Industria Manufacturera sin refinería, la venta de automóviles 0 km, la venta de nafta, y variables que informan sobre la actividad de Argentina y Brasil. Finalmente, cuando se analiza el desempeño de los modelos en las crisis, un modelo basado en EP resulta el más preciso en ambos casos.

Palabras claves: *Nowcasting; Ecuaciones Puente; Modelos de Factores Dinámicos; PIB; Uruguay.*

TABLA DE CONTENIDO

I.INTRODUCCIÓN.....	1
II. ANTECEDENTES	4
III. DISEÑO METODOLÓGICO	10
III.1 Modelos.....	10
III.1.1 Modelos de Factores Dinámicos	10
III.1.2 Modelos basados en Ecuaciones Puente	13
III.2 Preselección de variables y descripción de los datos	17
III.3 Estrategia empírica	19
III.3.1 Estimación de los modelos de Factores Dinámicos	21
III.3.2 Estimación de los modelos basados en Ecuaciones Puente	21
III.3.3 Estimación de los modelos <i>benchmark</i>	23
IV. RESULTADOS.....	24
IV.1 Desempeño de los distintos modelos	24
IV.2 Contribución de cada variable a la actualización de los pronósticos	29
IV.3 Pronósticos en momentos de grandes quiebres	34
IV.3.1 Crisis del 2002.....	35
IV.3.2 Crisis de 2020.....	40
V. SÍNTESIS DE RESULTADOS.....	47
VI. CONSIDERACIONES FINALES	49
BIBLIOGRAFÍA.....	51
ANEXO 1	57
ANEXO 2.....	61

I.INTRODUCCIÓN

Conocer el estado de situación de la actividad económica en tiempo real constituye un insumo fundamental para la toma de decisiones, tanto para los agentes económicos en general como para los hacedores de política. Sin embargo, las estadísticas oficiales que informan sobre las variables macroeconómicas de mayor interés presentan un considerable rezago respecto al período de referencia. El caso paradigmático es el de las estadísticas de contabilidad nacional, cuyas series, incluida la de Producto Interno Bruto (PIB), se suelen publicar internacionalmente con un rezago mayor a un mes y medio luego de finalizado el período de referencia. En Uruguay, el Banco Central del Uruguay (BCU), organismo encargado de elaborar las Cuentas Nacionales, publica las estimaciones de la actividad económica con frecuencia trimestral y aproximadamente a 80 días de finalizado el trimestre al que refieren. Por este motivo, las mismas no resultan oportunas para la toma de decisiones en tiempo real.

Este retraso en la disponibilidad de información adquiere mayor relevancia en los períodos en los que la evolución de la actividad económica presenta quiebres bruscos e inesperados, tal como el que se habría originado por el Covid-19. Cuando esto ocurre, captar el impacto del fenómeno a tiempo contribuye a que los agentes tomen sus decisiones con mayor y mejor información. En particular, en los períodos de crisis, si se toman medidas basándose en estimaciones erróneas del estado de situación de la economía, puede que el efecto de la recesión se amplifique o que la salida tenga plazos y costos mayores.

Por otra parte, principalmente debido a los avances tecnológicos, existe un mayor número de indicadores de alta frecuencia -mensual, semanal o diaria- que podrían aproximar el desempeño de la actividad y captar fluctuaciones en tiempo real, o al menos con un rezago menor al de las estadísticas de Cuentas Nacionales. Sin embargo, debido en parte a su alta frecuencia, y a que a diferencia de las cifras

oficiales de contabilidad nacional no suele estar cotejada su consistencia respecto a otras fuentes, el usar estos indicadores aisladamente podría brindar señales incorrectas e informar de forma parcial o divergente respecto a la información más confiable que se publica con mayor rezago.

De esta forma, se cuenta por un lado con información confiable pero menos oportuna, proveniente de las Cuentas Nacionales y, por otro lado, con información de alta frecuencia, pero más heterogénea, desorganizada y menos confiable. En este marco, dada la necesidad cada vez mayor de tener información precisa y en tiempo real, es que en las últimas dos décadas se ha desarrollado una importante literatura sobre lo que se denomina *Nowcasting*.

El término *Nowcasting* -que se origina de la unión de las palabras *now* (ahora) y *forecasting* (pronóstico)- es definido como la predicción del presente, el futuro muy cercano y el pasado muy reciente (Banbura et al., 2010). El principio del *Nowcasting* es la explotación de aquella información disponible antes y con mayor frecuencia que la variable objetivo (Banbura et al., 2013). Dicha información puede contener todo tipo de indicadores, desde encuestas de expectativas a agentes, variables financieras, encuestas a partir de muestras probabilísticas, registros administrativos, información que surge de la web, entre otros. Las técnicas de *Nowcasting* se basan en métodos puramente estadísticos que modelizan la relación de variables construidas en distintas frecuencias, con el fin de optimizar el uso de la información disponible en cada momento a los efectos de la predicción en tiempo real, no siendo un objetivo establecer ni interpretar relaciones causales.

Estos modelos de *Nowcasting*, al utilizar variables de alta frecuencia, en teoría, resultan útiles para predecir la actividad económica en momentos en los que la economía sufre cambios abruptos. No obstante, si las variables de alta frecuencia empleadas en la modelización no tienen la capacidad de captar correctamente los quiebres sufridos por la variable a pronosticar, es esperable que las estimaciones se comporten de forma errática. A pesar de ello, la literatura del *Nowcasting* no ha puesto

especial énfasis en el análisis de las estimaciones en los momentos en que existen estos quiebres. Dado que estos acontecimientos constituyen una de las principales fuentes de error en los pronósticos económicos (Hendry, 2006), resulta de interés analizar la performance predictiva de los modelos de *Nowcasting* en contextos de quiebre.

En este marco, el primer objetivo de este trabajo consiste en plantear modelos de *Nowcasting* para la predicción del PIB trimestral uruguayo. Concretamente, se estiman modelos de Factores Dinámicos y modelos basados en Ecuaciones Puente, con metodologías de estimación y series de datos novedosas respecto a los antecedentes nacionales. Un segundo objetivo radica en analizar la contribución de cada una de las variables utilizadas en las estimaciones a la actualización de los pronósticos. Finalmente, el tercer objetivo consiste en evaluar la capacidad predictiva de los modelos planteados en los últimos dos eventos en los que la actividad económica sufre fluctuaciones abruptas: la crisis del 2002 y la de 2020.

El resto del documento se organiza de la siguiente forma: en el capítulo II se describen los antecedentes bibliográficos. En el capítulo III se presenta el diseño metodológico: los modelos, los datos utilizados, y la estrategia empírica para la estimación y evaluación de los modelos. En el capítulo IV se desarrollan los resultados asociados a cada uno de los objetivos. En el capítulo V se realiza una síntesis de los principales hallazgos. Finalmente, en el capítulo VI se detallan las principales contribuciones empíricas y metodológicas, y posibles líneas de trabajo a futuro. Este documento se acompaña de dos anexos.

II. ANTECEDENTES

En las últimas dos décadas, sobre todo luego de la crisis de 2008, el interés sobre estimaciones en tiempo real para variables macroeconómicas se ha vuelto cada vez mayor, traduciéndose en un importante desarrollo de la literatura sobre *Nowcasting*.

Los enfoques metodológicos para el tratamiento de datos de distinta frecuencia pueden dividirse en cuatro grupos: (i) los modelos basados en Ecuaciones Puente (EP), (ii) los modelos de Factores Dinámicos (FD), (iii) los Modelos en Frecuencia Mixta (MIDAS), y (iv) los Vectores Autorregresivos en Frecuencia Mixta (FM-VAR).

A continuación, se presentan los principales antecedentes internacionales para las dos metodologías utilizadas en este trabajo: modelos basados en EP y modelos FD. Posteriormente, se desarrollan los antecedentes nacionales en la materia.

Los modelos basados en EP han sido de los primeros utilizados para el *Nowcasting*. Consisten en modelos uniecuacionales que permiten obtener estimaciones de las variables en baja frecuencia mediante la agregación de los indicadores de alta frecuencia (Forni y Marcellino, 2013).

La primera aplicación del método para *Nowcasting* tiene origen en Ingenito y Trehan (1996) donde, a partir de múltiples indicadores mensuales, se construyen EP con distintas especificaciones para predecir el nivel de actividad con frecuencia trimestral de Estados Unidos. Se encuentra que los modelos con pocas variables pueden ser mejores prediciendo que aquellos que incluyen más, y que el desempeño predictivo mejora a medida que aumenta la disponibilidad de datos mensuales.

Por su parte, Baffigi et al. (2004) proponen EP para el producto agregado y sus componentes por el enfoque del gasto para la Zona Euro. Los resultados dan cuenta que estos modelos tienen una mejor performance que la de los modelos *benchmark*¹

¹ En la mayoría de los trabajos en los que se evalúa la capacidad predictiva de los modelos de *Nowcasting*, se compara la predicción del modelo construido respecto a modelos *benchmark*, que suelen ser modelos autorregresivos, paseos aleatorios o algún otro tipo de modelo con los que es usual

no solo para el trimestre corriente sino también para el siguiente. Más tarde, Hahn y Skudelny (2008) complementan el trabajo anterior utilizando los componentes del PIB por el enfoque de la producción, y también encuentran que los modelos EP tienen una performance superior a los utilizados como *benchmark*.

Otro antecedente relevante para el presente trabajo es el de Diron (2008), cuyo objetivo es indagar si los ejercicios de proyección en el corto plazo realizados mediante ejercicios en *pseudo* tiempo real² arrojan distintas conclusiones respecto a los realizados en tiempo real. La autora muestra que en ciertos períodos hay diferencias en los errores de predicción según el tipo de ejercicio. No obstante, en promedio los errores de predicción son similares entre métodos, e inferiores a los *benchmark*.

Por su parte, los modelos de Factores Dinámicos son utilizados para predecir en tiempo real, aplicando como estrategia la sintetización de información de un amplio conjunto de variables en un número reducido de factores comunes.

Los trabajos de Stock y Watson (1998, 2002) son pioneros en la literatura sobre el uso de factores dinámicos para frecuencias mixtas. Estos autores proponen una metodología para predecir seis agregados macroeconómicos para Estados Unidos a partir de 215 variables, explotando mediante la estimación de factores, las dinámicas comunes existentes dentro de las series analizadas. Estos factores son estimados a partir del método de componentes principales. Pese a que en los desarrollos empíricos utilizan solo variables en frecuencia mensual, presentan un método teórico para trabajar con frecuencias mixtas.

Si bien se había demostrado la utilidad de la herramienta mediante aplicaciones prácticas, no es hasta Giannone et al. (2008) que se establece un marco de referencia formal para la predicción en tiempo real de variables en frecuencia trimestral mediante

predecir en el corto plazo. Para establecer si un modelo tiene mejor capacidad predictiva respecto a otro se contrastan los Errores Cuadráticos Medios de las predicciones.

² Los ejercicios en *pseudo* tiempo real consisten en evaluar predicciones fuera de la muestra con la última versión de los datos. En tanto, los ejercicios en tiempo real utilizan los datos que hubiesen estado disponibles al momento de predecir.

FD usando datos mensuales cuya actualización es no sincrónica³. Los autores plantean el modelo FD mediante una representación de tipo Estado Espacio, utilizando el filtro de Kalman para completar los datos faltantes de las variables mensuales⁴.

Con el mismo objetivo, Bańbura y Modugno (2014) formalizan una variante del algoritmo *Expectation Maximization* (EM) para, mediante la misma representación de tipo Espacio Estado de Giannone et al. (2008), obtener estimaciones máximo verosímiles de las variables y parámetros de interés. La variante propuesta consiste en modificar el EM propuesto por Watson y Engle (1983) y Shumway y Stoffer (1982) para poder plantear modelos en frecuencia mixta con un patrón de actualización de los datos no sincrónico. Doz et al. (2011) y Doz et al. (2012) demuestran la consistencia de los estimadores antes presentados.

Por último, dentro de estos modelos, un antecedente fundamental para el presente trabajo es el de Cascaldi-Garcia et al. (2021), quienes estiman un modelo FD para predecir la actividad de la Zona Euro y sus tres principales economías. Para ello, se utilizan como predictoras tanto variables mensuales específicas de cada país, como variables del resto de los países. Se demuestra que la inclusión de variables externas resulta importante a la hora de mejorar la capacidad predictiva de la tasa de crecimiento del PIB para las tres economías. A su vez, al analizar la secuencia de *nowcasts* antes y durante las últimas tres últimas recesiones que ha atravesado la Zona Euro, se encuentra que estos modelos tienen una relativa buena performance anticipando recesiones, ya que sus proyecciones resultan más acertadas que las de agentes privados.

³ La no sincronidad hace referencia a que los indicadores mensuales presentan distinto rezago de publicación, por lo que cuando se estiman modelos el último dato disponible de cada serie puede no referir al mismo mes.

⁴ Según lo establecen Doz et al. (2011) este método consta de dos etapas: una primera en la que se estiman factores mediante componentes principales, y una segunda donde se obtienen reestimaciones de los factores y de los parámetros mediante el filtro del Kalman.

Para el caso de Uruguay, se encuentran hasta el momento tres trabajos que buscan conocer la evolución de la actividad en tiempo real empleando alguna de las cuatro metodologías mencionadas: Masoller (2002), Rodríguez (2014) y Brum y Rodríguez (2016).

Masoller (2002) desarrolla un Indicador Sintético para la Actividad Económica uruguaya (ISAE), mediante la estimación de un modelo de Factores Dinámicos. El ISAE se compone de cuatro variables: la recaudación de IVA, el índice de volumen físico de la industria manufacturera, las importaciones de bienes sin petróleo, y las ventas de cemento a obras privadas. En el período de estimación (que abarca desde el primer trimestre de 1989 al segundo de 2001) el ISAE presenta una dinámica muy similar a la del PIB trimestral. Para evaluar la capacidad de predicción del PIB trimestral utilizando este indicador, se realizan predicciones *pseudo* fuera de la muestra, mediante la incorporación del ISAE trimestralizado en un sistema VAR que contiene el PIB trimestral. Los resultados muestran que la utilización del ISAE reduce los errores de predicción aproximadamente a la mitad respecto a los de un modelo autorregresivo.

Rodríguez (2014) estima modelos de Factores Dinámicos con el objetivo de obtener predicciones en tiempo real de la tasa de crecimiento del PIB trimestral a partir de variables en frecuencia mensual. El período de análisis abarca desde el primer trimestre de 1997 al primero de 2013. Se evalúa la capacidad predictiva de distintas especificaciones que combinan nueve variables, comparándolas entre sí y respecto a modelos *benchmark*, que en este caso fueron un AR(1) y un paseo aleatorio. Dentro de las diferentes especificaciones, el modelo con mejor performance incluye los siguientes regresores: el índice de volumen físico de la industria manufacturera, las ventas autos 0km, la tasa de empleo, las exportaciones e importaciones de bienes a precios constantes, las ventas de cemento y portland con destino mercado interno, el índice de confianza al consumidor y la cantidad de turistas. Los resultados obtenidos se encuentran en línea con la literatura: la capacidad predictiva del modelo FD supera

la de los modelos *benchmark*, no solo en la estimación del trimestre en curso, sino que también en las predicciones del siguiente.

Brum y Rodríguez (2016), con el fin de predecir el estado actual de la economía en el corto plazo, utilizan la metodología de modelos basados en Ecuaciones Puente. Igual que Rodríguez (2014), la variable a predecir es la variación del PIB con frecuencia trimestral, utilizando como variables explicativas indicadores en frecuencia mensual. Se estiman modelos mediante dos métodos, que los autores denominan directo e indirecto. En el directo, el PIB trimestral es estimado a partir de tres variables: ventas de autos 0km, un índice de volumen físico del valor agregado de la industria manufacturera, y exportaciones de bienes a precios constantes. Mientras que en el indirecto, para cada componente del PIB por el enfoque de la producción -esto es, para cada uno de los valores agregados (VAB) sectoriales- se estima una ecuación puente que relaciona dicha variable en frecuencia trimestral con la misma variable, pero en frecuencia mensual. Este dato surge a partir del Índice Mensual de Actividad Económica (IMAE) de uso interno en el BCU. Para este modelo indirecto, la predicción del PIB se obtiene agregando las predicciones de los distintos VAB sectoriales. Los resultados obtenidos muestran que el modelo indirecto tiene una mejor capacidad predictiva del estado de la economía, tanto respecto a los modelos *benchmark*, como al directo, modelo que a su vez cuenta con mejor capacidad predictiva que los *benchmark*. Asimismo, se evalúan estos modelos respecto al de Rodríguez (2014), resultando que este último tiene mejor performance que el directo, y similar al indirecto.

En base al estado de situación de la literatura uruguaya, en este trabajo se estiman modelos con las metodologías EP y FD, con la particularidad de que en los modelos EP las variables son seleccionadas recursivamente durante el período de estimación a través de los métodos GETS y LASSO. Por su parte, se analiza la contribución de las variables incluidas en los modelos a las actualizaciones de los pronósticos, lo que

también constituye una novedad en el contexto nacional. Asimismo, se incorpora el estudio del desempeño de los modelos en las crisis de 2002 y 2020.

III. DISEÑO METODOLÓGICO

Este capítulo se organiza en tres secciones. En la primera se desarrollan las metodologías empleadas para estimar los modelos. En la segunda se presentan las series utilizadas para las estimaciones. En tercer lugar, se explicita la estrategia empírica llevada a cabo para la estimación y evaluación de los modelos.

III.1 Modelos

A continuación, se presentan los modelos de Factores Dinámicos (FD) y los basados en Ecuaciones Puente (EP), detallando en cada caso particularidades asociadas a transformaciones necesarias en las variables, métodos de proyección, y las metodologías utilizadas para la elección de las variables que forman parte de las estimaciones.

III.1.1 Modelos de Factores Dinámicos

Los modelos que utilizan factores sintetizan la información de un amplio conjunto de variables en un reducido número de factores. Al estimar estos factores, se logran modelos parsimoniosos capaces de conservar aquellas características de las variables que obedecen a dinámicas comunes. Un supuesto fundamental consiste en considerar a cada variable como la suma de dos componentes ortogonales: el primero, que es compartido con el resto de las variables que forman parte de los factores comunes y está asociado a co-movimientos entre las mismas, mientras que el segundo es considerado un componente idiosincrático, propio de cada variable e independiente del resto (Sargent y Sims, 1977; Giannone et al., 2005) .

Los modelos de Factores Dinámicos (FD) para variables en frecuencia mensual pueden resumirse en las siguientes ecuaciones:

$$y_t = \gamma + \Lambda f_t + e_t \quad (1)$$

$$f_t = \sum_{j=1}^P A_j f_{t-j} + u_t, \quad u_t \sim i.i.d. N(0, Q), \quad (2)$$

donde $y_t = (y_{1,t}, y_{1,t}, \dots, y_{N,t})'$ es un vector de dimensión $N \times 1$ con variables en frecuencia mensual y estacionarias, para los meses $t = 1, 2, \dots, T$; γ es un vector de N constantes, $f_t = (f_{1,t}, f_{2,t}, \dots, f_{r,t})'$ de dimensión $r \times 1$ que contiene r factores inobservables; $\Lambda = (\lambda'_{1,t}, \lambda'_{2,t}, \dots, \lambda'_{N,t})'$ una matriz de dimensión $N \times r$, con coeficientes que relacionan los r factores inobservables con las N variables incluidas en el modelo.

Para completar la caracterización de (1), $e_t = (e_{1,t}, e_{2,t}, \dots, e_{N,t})'$ es un vector de $N \times 1$ shocks idiosincráticos, que siguen un proceso AR(1):

$$e_{i,t} = \alpha_i e_{i,t-1}, \quad e_{i,t} \sim i.i.d. N(0, \sigma_i^2) \quad \forall i = 1, 2, \dots, N, \quad (3)$$

con $E[e_{i,t}, e_{j,s}] = 0$ cuando $i \neq j$.

El término Λf_t es el componente común y se encarga de modelar los co-movimientos entre las N variables incluidas en el modelo. Esto modela la supuesta existencia de shocks comunes que afectan a todas las variables, pero con distinta magnitud, la que está dada por los valores de Λ . Mientras que e_t representa aquellos shocks que repercuten individualmente en cada variable i , independientes a los del resto de las $(N - 1)$ variables.

En la ecuación (2), A_1, A_2, \dots, A_P son matrices de dimensión $r \times r$ con los coeficientes autorregresivos asociados a cada factor.

III.1.1.1 Frecuencia mixta

El modelo presentado anteriormente se encuentra especificado en frecuencia mensual. Debido a que la variable de interés es el PIB trimestral, resulta necesario incorporarla a la modelización. En la literatura de los modelos FD es usual hacerlo mediante la transformación propuesta por Mariano y Murasawa (2003). Estos autores

derivan una aproximación en la que se considera a la variable trimestral como una variable mensual con datos faltantes.

La transformación tiene como resultado la siguiente expresión, en la que la tasa de crecimiento del PIB trimestral es expresada en función de las tasas de crecimiento del PIB en frecuencia mensual (variable inobservable):

$$y_t^Q = \frac{1}{3}(y_t^M + 2y_{t-1}^M + 3y_{t-2}^M + 2y_{t-3}^M + y_{t-4}^M), \quad t = 3,6,9,12, \dots, 3T, \quad (4)$$

donde y_t^Q es la tasa de crecimiento del PIB trimestral aproximada por la diferencia logarítmica, mientras que y_t^M es la tasa de crecimiento del PIB mensual, aproximada también por la diferencia logarítmica.

III.1.1.2 Representación Estado Espacio

La metodología seguida en este trabajo para estimar los componentes inobservables y posteriormente hacer predicciones de la variable de interés es la propuesta por Bańbura y Modugno (2014), que consiste en una adaptación del algoritmo *Expectation Maximization* (EM) (Dempster et al., 1977; Watson y Engle, 1983), para mediante la aplicación del filtro de Kalman, obtener estimaciones cuando existen observaciones faltantes.

Para poder aplicar EM en un modelo FD con la transformación presentada anteriormente, éste debe ser planteado mediante una representación del tipo Estado Espacio (EE). Las representaciones EE se conforman de dos ecuaciones: la de Medida, que describe como las variables observadas dependen de un estado inobservable, y la de Estado, que especifica la dinámica de dicho estado inobservable.

$$x_t = \tilde{\Lambda}s_t + \eta_t, \quad (\text{Ecuación de Medida})$$

$$s_t = \tilde{A}s_{t-1} + v_t \quad (\text{Ecuación de Estado})$$

En la Ecuación de Estado, s_t es un vector de variables inobservables, \tilde{A} se conoce como la matriz de transición y contiene los parámetros asociados a la dinámica de las variables inobservables, mientras que η_t es un término de error. En la Ecuación de

Medida, $\tilde{\Lambda}$ es una matriz que contiene los coeficientes, que asocia inobservables con x_t , vector que contiene las variables observadas en ambas frecuencias, al tiempo que v_t es un término de error. En el Anexo 2 se detalla de forma matricial cada uno de los componentes de ambas ecuaciones.

Los parámetros del modelo, así como las observaciones faltantes e inobservables, se estiman mediante máxima verosimilitud, planteando la función de verosimilitud como si la información fuese completa, y mediante sucesivas iteraciones -utilizando el filtro de Kalman- converger a los valores óptimos (Bańbura y Modugno, 2014).

Para el funcionamiento del algoritmo es necesario fijar valores iniciales para las variables y parámetros. Para esto se sigue a Bańbura y Modugno (2014), que aplican la primera etapa del método desarrollado por Giannone et al. (2008), consistente en estimar los factores y los coeficientes de las matrices \tilde{A} mediante componentes principales, restringiendo la muestra a un panel balanceado de datos.

III.1.2 Modelos basados en Ecuaciones Puente

Los modelos basados en Ecuaciones Puentes (EP) constituyen la primera técnica econométrica desarrollada para trabajar con variables en frecuencia mixta. La técnica consiste en utilizar regresiones lineales para relacionar una variable de baja frecuencia con un reducido número de variables en una frecuencia mayor, a través de su correcta agregación. De esta forma, los modelos quedan especificados en la menor de las frecuencias, en este caso la trimestral.

Si bien el uso de EP es muy variado en cuanto a la especificación del modelo, la forma más general puede expresarse como:

$$y_t^Q = \alpha + \sum_{i=1}^P \beta_i y_{t-3i}^Q + \sum_{j=1}^N \sum_{k=0}^S \gamma_{j,k} x_{j,t-3k}^Q + e_t, \quad (5)$$

siendo y_t^Q la variable de interés con frecuencia trimestral para los meses $t = 3, 6, 9, 12, \dots, 3T$; α una constante; β_i , P parámetros con $i = 1, 2, \dots, P$ (P representa el

último rezago de la variable de interés incluida en el modelo), x_j^Q , el j -ésimo indicador mensual agregado de forma que represente correctamente su contraparte trimestral⁵ y $\gamma_{j,k}$, $(S + 1) * N$ parámetros asociados a los N indicadores mensuales incluidos (tanto para sus S rezagos, como para los valores contemporáneos). En el caso que no se disponga de las variables mensuales hasta el período necesario, las mismas deben ser proyectadas.

De esta forma, para especificar modelos basados en EP hay que definir qué variables mensuales incluir, cuántos rezagos de la variable dependiente y de las independientes, y cómo proyectar las variables mensuales para completar las observaciones faltantes.

La proyección de las series se resuelve a través de la función *autoarima*. Dicha función es una variación del algoritmo Hyndman-Khandakar⁶, que combina pruebas de raíz unitaria, minimización del criterio Akaike (AIC) y estimaciones por máxima verosimilitud a los efectos de seleccionar el modelo ARIMA que mejor ajusta para cierta serie temporal.

El algoritmo de selección determina, en primer lugar, el número de diferencias entre 0 y 2 mediante la prueba de raíz unitaria KPSS⁷. En segundo lugar, selecciona la estructura autorregresiva (p) y de medias móviles (q), mediante AIC. Una vez que se encuentra el modelo con menor AIC, se realizan pruebas sumando y restando en 1 a p y q , y agregando y quitando la constante, hasta obtener el modelo que efectivamente minimiza el criterio AIC.

En las siguientes dos secciones se explica cuáles son para este trabajo los métodos empleados para abordar las decisiones referidas a la inclusión de variables y rezagos:

⁵ Para las variables de stock, se considera el valor del último mes del trimestre, mientras que para las variables de flujo se toma la suma de los valores de los meses del trimestre si la variable se encuentra en nivel o el promedio para las expresadas en forma de índice con base fija.

⁶ Véase Hyndman y Khandakar (2008).

⁷ Véase Kwiatkowski et al. (1992).

General to Specific (GETS) y Least Absolute Shrinkage and Selection Operator (LASSO).

III.1.2.1 Método GETS

El método GETS tiene como principal objetivo la obtención de modelos econométricos parsimoniosos, en los que se busca relacionar eficientemente la evidencia empírica con la teoría económica (Campos et al., 2005). Existen numerosas contribuciones que han probado la efectividad del método, y han avanzado en desarrollos teóricos e implementaciones informáticas (Hoover y Pérez, 1999; Hendry et al., 2007; Doornik y Hendry, 2007; Doornik, 2009).

GETS resulta útil en los casos en los que se quiere especificar un modelo a partir de un conjunto de variables candidatas:

$$y_t = \alpha + \sum_{i=1}^P \beta_i y_{t-3i} + \sum_{j=1}^N \sum_{k=0}^S \gamma_{j,k} x_{j,t-3k} + e_t, \quad (6)$$

donde y_t es la variable dependiente que se quiere modelar en los meses $t = 3, 6, 9, 12 \dots, 3T$ mediante relaciones con su propio pasado, así como con valores contemporáneos y pasados de un conjunto de N variables x_i con $i = 1, 2, \dots, N$. Mientras que e_t es un término de error con distribución normal.

El método asume que existe al menos un proceso generador de datos local (PGDL) anidado en la ecuación (6). Esto significa que en dicha ecuación existe al menos una especificación que es estadísticamente válida a los efectos de representar el proceso generador de los datos (PGD) de y_t (Hendry y Doornik, 2014).

El proceso de selección de variables mediante GETS es iterativo. Se inicia planteando un modelo general (GUM) con todas las variables y rezagos candidatos a ser seleccionados. Una vez chequeada la normalidad y no autocorrelación de los errores del GUM se procede al segundo paso, donde se elimina una a una aquellas variables no significativas, generándose así diversos nodos. Cada uno de dichos nodos, se

compone de las variables del GUM exceptuando de a una las variables no significativas del paso previo. El proceso se repite sucesivas veces, generándose nuevos nodos a partir de los generados en la etapa anterior. En cada etapa, se validan los nodos a través de criterios de información y pruebas que tienen en cuenta entre otras cosas la parsimonia. Cuando se llega al fin del proceso de eliminación de variables se selecciona el modelo más adecuado de acuerdo con algún criterio de información (Doornik, 2019).

III.1.2.2 Método LASSO

El método LASSO, introducido por Tibshirani (1996) ha ganado popularidad en la última década dentro de la literatura de *machine learning*. El método es similar a MCO, en cuanto a que la función de pérdidas que se optimiza para obtener los parámetros es la misma, con la diferencia de que se agrega una función de penalización extra, que incorpora la posibilidad de que los parámetros asociados a algunos regresores sean cero. Por lo que la función a minimizar para obtener los coeficientes mediante LASSO es la siguiente:

$$\sum_{t=1}^T (y_t - \mu - \sum_{i=1}^K z_{t,i} \delta_i)^2 + \lambda \sum_{i=1}^k |\delta_i|, \quad (7)$$

donde y_t es la variable de interés para $t = 3, 6, 9, 12 \dots, 3T$, μ es el intercepto, $z_{t,i}$ es un conjunto de K variables con $i = 1, 2, \dots, K$ que contiene rezagos de y_t y un conjunto de regresores contemporáneos y rezagados respecto a t . Por último λ es un parámetro que penaliza la cantidad de variables que efectivamente forman parte del modelo, a mayor valor del parámetro λ , más parámetros toman el valor cero. Nótese que en el caso extremo de que $\lambda = 0$ se estaría en el caso de MCO donde todas las variables forman parte del modelo, al tiempo que si $\lambda = \infty$ el valor óptimo para los $K \delta_i$ es cero.

Para determinar los valores de λ se sigue lo sugerido por Belloni y Chernozhukov (2011).⁸

III.2 Preselección de variables y descripción de los datos

En esta sección se presentan las variables utilizadas en la investigación, fundamentando su inclusión en los modelos. Se consideran series en dos frecuencias: trimestral, únicamente para la variable a predecir, y mensual, para todas las variables independientes con frecuencia mayor a la trimestral. A pesar de la existencia de variables con mayor frecuencia que la mensual, siguiendo a Bańbura et al. (2013) se incluyen solo mensuales. Estos autores muestran que el uso de variables con frecuencia semanal y diaria no mejora la capacidad predictiva de los modelos principalmente debido a su alta volatilidad, tanto en momentos ordinarios como de crisis.

La variable de interés a predecir es la tasa de crecimiento interanual del PIB trimestral uruguayo a precios constantes de 2016. El período de estimación del presente trabajo es 1994-2021, buscando así capturar los potenciales quiebres de nivel ocurridos en los años 2002 y 2020. Dado que no existe una serie que abarque la totalidad del período, la misma es construida a partir de tres fuentes: desde 2016 se utiliza la serie de Cuentas Nacionales (CCNN) base 2016, publicada por el BCU, que inicia en dicho año; para el período 1997-2015 se utiliza la serie que surge de Álvarez et al. (2021); mientras que para 1994-1996 se prolonga la serie 1997-2015 a través del método de tasas de variación, usando como indicadora la serie de CCNN fuente BCU base 1983.

Para la elección de las variables con frecuencia mensual, se tiene en cuenta que sean variables que desde la teoría económica y/o la lógica de las identidades de CCNN

⁸ Belloni y Chernozhukov (2011) establecen $\lambda = \frac{2\sigma\sqrt{U_t}}{\Phi(1-\tau/2K)}$, siendo σ la raíz cuadrada de la varianza del modelo con las K variables candidatas, U_t el número de períodos de las series, tal que si el subíndice t indica meses: $U_t = t/3$, Φ es la función de distribución acumulada de una normal estándar para cierto nivel de significación τ .

tengan potencialmente buen desempeño prediciendo el PIB uruguayo y, adicionalmente, que su publicación sea oportuna (considerando variables con un rezago menor a los 60 días). A su vez, es necesario que las series cuenten con suficientes observaciones para predecir el año 2002. En Uruguay existen relativamente pocas series que cumplen con dicho requisito. No obstante, en la literatura se muestra que utilizar un mayor número de variables, no implica necesariamente mejoras en la capacidad predictiva de los modelos (Alvarez et al., 2012; Bańbura y Modugno, 2014).

Como punto de partida para la selección de variables, se tienen en cuenta ocho de las utilizadas en Rodríguez (2014)⁹, ya que allí se sigue criterios similares a los propuestos en este trabajo y se demuestra su efectividad a la hora de predecir la dinámica del PIB. Dichas variables son: el Índice Mensual de la Industria Manufacturera (*IVF_IM*), la tasa de empleo (*temp*), las ventas de automóviles 0km (*vautos*), las importaciones (*impb*) y exportaciones (*expb*) de bienes en términos de volumen físico, la cantidad de turistas que ingresan al país (*turistas*), las ventas de combustibles (*gasoil* y *nafta*) y las ventas de cemento al mercado interno (*cemento*).¹⁰

Además, en este trabajo se incorporan nuevas variables teniendo en cuenta su potencial buen desempeño: el Índice de Salario Real (*isr*), la recaudación en términos reales del Impuesto al Valor Agregado (*iva*), las cabezas faenadas de ganado ovino y bovino (*faena*), el Indicador Mensual de Actividad Económica para Argentina (*EMAE*), el Índice de Actividad Económica del Banco Central de Brasil (*IBC-Br*), el Uruguay Bond Index (*UBI*) y el Índice Mensual de la Industria Manufacturera sin refinería (*IVF_IMsr*).

⁹ Las variables utilizadas por Brum y Rodríguez (2016) que son públicas, están contenidas en Rodríguez (2014)

¹⁰ La variable Índice de Confianza al Consumidor con fuente Equipos Mori - Universidad Católica, no fue considerada, ya que se encuentra disponible desde 2009.

Tanto la inclusión del Índice de Salario Real como de la recaudación de IVA se fundamenta en su esperada relación con el consumo de los hogares. En tanto, la incorporación de la faena responde a su correlación con el valor pecuario e industrial. Por su parte, los indicadores de actividad de Argentina y Brasil pretenden, por un lado, introducir de forma anticipada shocks con origen externo que pueden impactar en la economía nacional, no capturados o capturados con rezago por las variables internas disponibles. Por otro lado, buscan explotar la relación existente entre el nivel de actividad uruguayo y el de los dos países vecinos, explicada fundamentalmente por la importancia de las relaciones comerciales bilaterales, así como por la similitud en cuanto a los bienes que comercializan con el resto del mundo (Lanzilotta et al., 2003; (Dianessi et al., 2020).

La inclusión del UBI responde a que la literatura empírica suele incluir en los modelos variables financieras, que capturan rápidamente los cambios que ocurren en momentos de crisis o inestabilidad (el UBI es un ejemplo de ello).

Por último, al considerar el *IVF_IM* excluyendo la refinería se busca aislar la intermitencia de la actividad de esta última, debido a los cierres para mantenimiento.

En el Anexo 1 se presentan cada una de las variables, detallando su fuente, rezago y correlación con el PIB.

III.3 Estrategia empírica

La muestra empleada para estimar cada uno los modelos abarca desde el primer trimestre de 1994 hasta el mismo trimestre de 2021, generándose predicciones fuera de la muestra para los trimestres comprendidos entre el cuarto trimestre de 2001 y el segundo de 2021, período que se denomina “de evaluación”.

La estimación y evaluación de los modelos se hace mediante un ejercicio en *pseudo* tiempo real. Se considera como *pseudo* tiempo real y no tiempo real, debido a que, si bien se tuvo en cuenta para cada una de las variables su verdadero rezago de

publicación, se utiliza la última versión de los datos y no los que hubiesen estado disponibles en cada momento del pasado. De esta forma, se espera que las series utilizadas tengan un componente asociado a revisiones. No obstante, hay evidencia de que las conclusiones respecto a la bondad predictiva de los modelos no varía respecto a si se aplicara un ejercicio propiamente en tiempo real, ya sea si se estiman modelos a partir de EP (Diron, 2008) o modelos de FD (Schumacher y Breitung, 2008).

Para cada uno de los trimestres comprendidos dentro del período de evaluación 2001Q4-2021Q2 se realizan seis predicciones en *pseudo* tiempo real, separadas entre sí por 30 días. La primera predicción para un trimestre particular se realiza incorporando hasta la última observación de cada variable que hubiese estado disponible si se hubiese querido predecir el trimestre 180 días antes de conocer el dato oficial de CCNN. Las restantes cinco predicciones se realizan con la información que hubiese estado disponible 150, 120, 90, 60 y 30 días antes de la publicación. La fecha de actualización de los *nowcasts* se fija en el día 20 de cada mes, fecha aproximada en la que cada tres meses se conocen los datos del PIB trimestral. A modo de ejemplo, en la Tabla 1 se muestra la secuencia de los seis *nowcasts* para 2021Q2.

Tabla 1. Ejemplo de secuencia de predicciones para 2021Q2

Número de pronóstico para 2021Q2	Fecha de la predicción	Último dato trimestral disponible	Días antes de la publicación de 2021Q2	Horizonte de predicción
1	20-03-21	2020Q4	180	t+2
2	20-04-21	2020Q4	150	t+2
3	20-05-21	2020Q4	120	t+2
4	20-06-21	2021Q1	90	t+1
5	20-07-21	2021Q1	60	t+1
6	20-08-21	2021Q1	30	t+1

Nota: En los meses que se publican los datos de Cuentas Nacionales, las predicciones se hacen una vez conocidos los mismos. Para cada fecha de predicción, el mes al que refiere el último dato de cada una de las variables mensuales depende de su rezago.

Fuente: Elaboración propia.

La contraparte de esas seis estimaciones para un mismo trimestre es que en cada fecha de predicción se realizan en forma simultánea dos predicciones, una a un paso

(t+1) y otra, a dos (t+2). En la Tabla 2 se muestra, como ejemplo, la secuencia de pronósticos mensuales para el período comprendido entre marzo de 2021 y agosto de 2021.

Tabla 2. Ejemplo de secuencia de predicciones entre marzo y agosto de 2021

Horizonte de predicción	Fecha de la predicción					
	20-03-21	20-04-21	20-05-21	20-06-21	20-07-21	20-08-21
t+1	2021Q1	2021Q1	2021Q1	2021Q2	2021Q2	2021Q2
t+2	2021Q2 publicación de 2020Q4	2021Q2	2021Q2	2021Q3 publicación de 2021Q1	2021Q3	2021Q3

Nota: Se presenta para cada fecha de predicción comprendida entre marzo y agosto de 2021, los trimestres que se predicen a uno y dos pasos.

Fuente: Elaboración propia.

Todos los modelos se estiman con las variables en primeras diferencias del logaritmo de las series desestacionalizadas. La desestacionalización es implementada mediante el programa X-13 ARIMA, usando el paquete del *software* R documentado por Sax y Eddelbuettel (2018). A continuación, se describe el proceso de estimación y las decisiones adoptadas.

III.3.1 Estimación de los modelos de Factores Dinámicos

Para estimar el modelo de FD propuesto en este trabajo, con las variables que se detallan en el Anexo 1 (columna FD_A), la determinación del número r de factores a incluir se hace mediante el criterio de información BG, desarrollado por Bai y Ng (2002). Como resultado, el modelo se estima con un solo factor.

En tanto, el número p de rezagos para el factor común que modela su dinámica se determina mediante criterio Akaike. Se elige el número de rezagos que minimiza dicho criterio, a partir de una serie con un factor estimado mediante componentes principales, resultando en que dicha dinámica se especifica con un único rezago.

III.3.2 Estimación de los modelos basados en Ecuaciones Puente

Para la estimación de los modelos basados en EP se utilizan dos tipos de muestras denominadas: ventana expandible y ventana deslizante. En la ventana expandible, el

ejercicio en *pseudo* tiempo real es llevado a cabo mediante la incorporación de nuevas observaciones sin descartar las más antiguas, teniendo con el paso del tiempo muestras con un mayor número de observaciones en cada una de las series. En cambio, en ventana deslizante se estiman modelos con muestras de largo fijo, en este caso de 40 trimestres, descartándose sucesivamente las observaciones más antiguas.

Para proceder mediante GETS y LASSO se fija un valor máximo de ocho rezagos para la variable dependiente. Por su parte, para las agregaciones trimestrales de las variables mensuales se incluyen únicamente valores contemporáneos. Tal decisión, se justifica en que se realiza el ejercicio de estimar los modelos con hasta cuatro rezagos de las variables explicativas que da por resultado la selección de variables rezagadas en detrimento de las contemporáneas, yendo contra el principio del *Nowcasting* de utilizar la información más reciente y oportuna, y sin que esto cause mejoras considerables en el ajuste de los modelos.

Una de las novedades que se desarrollan en la presente investigación es que la selección de variables explicativas y rezagos del PIB trimestral se realiza de forma recursiva para todo el período, actualizándose cada vez que se cuenta con una nueva observación del PIB trimestral. En tanto, en la literatura de los modelos basados en EP es habitual que se seleccione mediante algún criterio un set de variables y rezagos que son mantenidos para todo el período de evaluación, mientras que se estiman de forma recursiva los parámetros.

En este trabajo tanto la selección de variables como la estimación de parámetros se actualiza cada tres meses. Siguiendo el ejemplo de la Tabla 2, los pronósticos hechos en los meses de marzo, abril y mayo de 2021, se realizan con los mismos parámetros y variables que surgen de las selecciones con LASSO y GETS con el último panel de datos balanceado de variables observadas, que en este caso es el que termina en 2020Q4. En tanto, los correspondientes a los meses de junio, julio y agosto de 2021, se hacen con el panel de datos que termina en 2021Q1.

En las estimaciones mediante LASSO, el parámetro λ que penaliza la inclusión de variables, también es estimado de forma recursiva. Por su parte, en las estimaciones mediante GETS se utiliza de forma recursiva la función Impulso Saturación,¹¹ que permite detectar automáticamente observaciones atípicas en la variable dependiente, de modo que cuando un atípico es identificado, el mismo se incorpora en la parametrización.

Esta parametrización recursiva permite mayor flexibilidad en los modelos, de forma que éstos puedan adaptarse a la evolución de las dinámicas relativas entre las variables.

III.3.3 Estimación de los modelos *benchmark*

Como es usual en la literatura empírica de *Nowcasting*, en el presente trabajo se evalúa la capacidad predictiva de los modelos en relación a modelos univariados, modelos que suelen tener un buen desempeño predictivo en el corto plazo. La estimación de dichos modelos *benchmark* también se realiza mediante ejercicios en *pseudo* tiempo real.

Los modelos *benchmark* estimados son cuatro: un autorregresivo de orden 1 AR(1), un modelo *autoarima*, un modelo en el que los parámetros y rezagos se determinan con GETS, y un modelo en los que los parámetros y rezagos se determinan mediante LASSO. Tanto los parámetros, como los rezagos -cuando corresponde-, se estiman de forma recursiva.

¹¹ Por detalles, véase Hendry et al. (2007); y Doornik (2009).

IV. RESULTADOS

Los resultados se agrupan en tres secciones. En la primera, se analiza el desempeño predictivo de los modelos propuestos para el período de evaluación. En la segunda, se pone el foco en la incidencia de cada una de las variables en las actualizaciones de los pronósticos. En tanto, en la tercera, se estudia la capacidad predictiva en los años 2002 y 2020.

IV.1 Desempeño de los distintos modelos

La evaluación de la capacidad predictiva de los modelos se realiza computando los errores de predicción *pseudo* fuera de la muestra, para las predicciones que estos realizan de la tasa de crecimiento interanual del PIB trimestral.

El estadístico utilizado para resumir la capacidad predictiva es la raíz del error cuadrático medio del error de pronóstico (RMSFE, según su sigla en inglés):

$$RMSFE = \sqrt{\frac{\sum_{t=1}^T (E[\hat{y}_{tr}^q | X_{v_r}] - y_t^q)^2}{T}}, \quad (8)$$

donde T es el número de observaciones; y_t^q el verdadero valor de la tasa de crecimiento interanual del PIB trimestral en t ; $E[\hat{y}_{tr}^q | X_{v_r}]$ es la predicción de y_t^q realizada en el momento r condicional a la información disponible, X_{v_r} , en r . En este caso, r es una variable mensual que representa el día 20 de cada mes, momento en los que se realizan los pronósticos, 180, 150, 120, 90, 60 y 30 días antes de conocer y_t^q .

En la Tabla 3 se presentan los sucesivos RMSFEs para los pronósticos de los siguientes modelos: modelo basado en EP con ventana expandible usando LASSO (Lasso_E), modelo basado en EP con ventana deslizante usando LASSO (Lasso_D), modelo basado en EP con ventana expandible usando GETS (Gets_E), modelo basado en EP con ventana deslizante usando GETS (Gets_D), el modelo FD

propuesto en este trabajo (FD_A), el modelo basado en EP directo propuesto por Brum y Rodríguez (2016) (EP_BR) y el mejor de los modelos de FD propuestos por Rodríguez (2014) FD_R¹².

Los resultados se presentan agrupados por el adelanto respecto a la publicación de los datos oficiales. Los valores de la tabla están expresados en porcentajes, con relación al modelo *benchmark* que presenta mejor desempeño fuera de la muestra, el *autoarima*. Los porcentajes informan cuánto más pequeños son los RMSFEs del modelo correspondiente, respecto a los del *autoarima*. En la tabla también se presentan los RMSFEs para el *autoarima*, en este caso en términos absolutos. En estos resultados se excluyen los errores de predicción para el tercer trimestre de 2002 y el segundo de 2020, debido a que en estos dos trimestres la magnitud de los errores es muy influyente en el estadístico, no permitiendo extraer conclusiones sobre la evolución de la capacidad predictiva para el resto del período de evaluación. El análisis para estos trimestres será presentado en la sección IV.3.

¹² Para estimar el modelo de Brum y Rodríguez (2016) se utiliza la serie de IVF_M fuente INE en lugar de la serie de uso interno BCU. Para el modelo de Rodríguez (2014) se presentan los resultados del modelo que incluye el IVF_IM, la tasa de empleo, las ventas de automotores, importaciones y exportaciones de bienes, cantidad de turistas, ventas de cemento y ventas de combustibles.

Tabla 3. Raíz del error cuadrático medio del error de pronóstico (RMSFE), en porcentajes

Modelo	Horizonte de predicción					
	Dos pasos (t+2)			Un paso (t+1)		
	Días antes de la publicación de los datos oficiales					
	180	150	120	90	60	30
Gets_E	26,9 ()	28,9 ()	38,4 (**)	39,1 (**)	41,0 (**)	41,2 (**)
Gets_D	13,9 ()	17,7 ()	20,8 ()	26,2 (*)	25,5 (*)	25,7 (*)
Lasso_E	30,1 ()	31,9 ()	38,1 (*)	39,1 (**)	42,0 (**)	42,4 (**)
Lasso_D	29,8 ()	35,1 ()	38,2 (**)	38,2 (**)	40,5 (**)	40,9 (**)
DF_A	27,9 ()	33,4 ()	37,0 (*)	33,9 (*)	37,3 (**)	41,3 (**)
EP_BR	17,6 ()	23,2 ()	28,5 ()	24,2 ()	28,2 (*)	31,6 (*)
FD_R	26,8 ()	31,1 ()	33,5 (*)	29,3 (*)	33,7 (*)	37,8 (**)
<i>autoarima</i>	3,10	3,10	3,10	1,96	1,96	1,96

*Nota: Para el modelo autoarima se presentan los RMSFEs absolutos. Para el resto de los modelos, los RMSFEs se presentan con relación al autoarima y se calculan para cada período de actualización como la diferencia entre los RMSFEs del benchmark y los de cada modelo, dividiéndose sobre los RMSFE del primero. Entre paréntesis se presentan los p-valores de las pruebas de Diebold-Mariano, ***p<0,01, **p<0,05, *p<0,1.*

Fuente: Elaboración propia.

La Tabla 3 muestra que para todos los modelos evaluados, la sola inclusión de las variables explicativas mejora la capacidad predictiva. Esto se infiere a partir de la columna a 30 días, donde los modelos son estimados sin datos mensuales proyectados, siendo para este caso los RMSFEs entre 25,7% y 42,4% más bajos que los del *benchmark*. Esta mejora se observa a su vez para los otros cinco períodos de actualización de los pronósticos en los que las variables explicativas mensuales son proyectadas. Entonces, una primera conclusión es que, aun cuando se hacen predicciones con observaciones faltantes -para algunas o todas las variables- la inclusión de variables mensuales hace que los errores de predicción sean inferiores que los del *benchmark*, justificándose en cierta forma la utilización de estos modelos para predecir en el corto plazo.

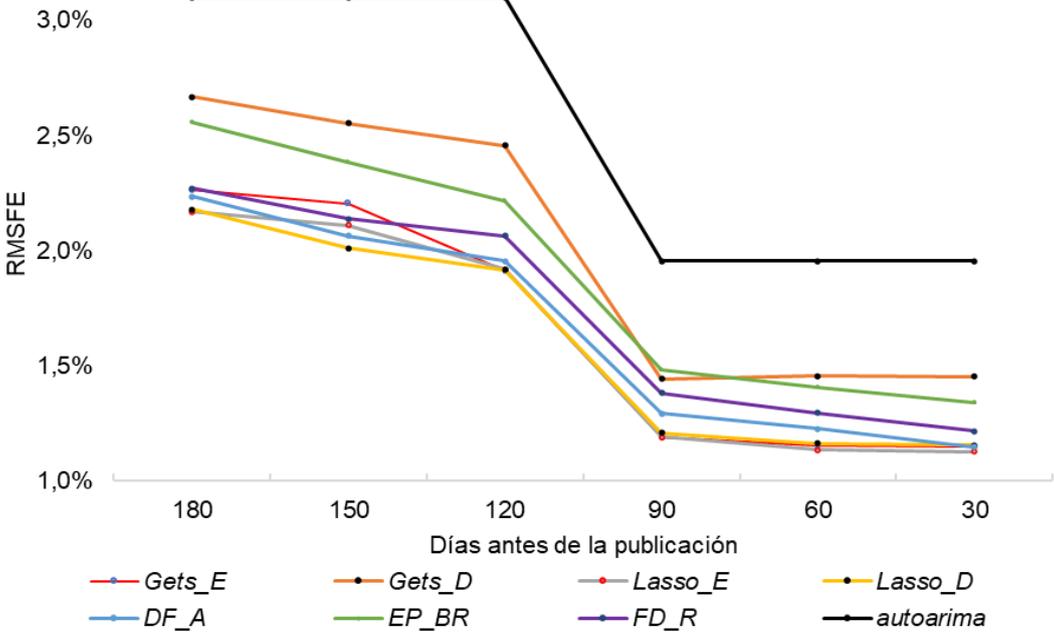
Para determinar si estas mejoras en los errores son estadísticamente significativas, se utiliza la prueba de igualdad de predicción DM, propuesta en Diebold y Mariano (1995).¹³ Los resultados muestran que para (t+1) seis de los siete modelos que

¹³ La hipótesis nula de la prueba es que los pronósticos de los dos modelos tienen la misma capacidad predictiva, mientras que la alternativa es que la capacidad predictiva del modelo a evaluar es mejor que la del *benchmark*.

incorporan variables mensuales del trimestre que se proyecta tienen errores inferiores que el *benchmark*, con un nivel de significación menor o igual al 10%. Por su parte, para (t+2) los resultados de la prueba son heterogéneos dependiendo del modelo y del momento del pronóstico. Un aspecto a destacar respecto a la prueba para (t+2) es que varios de los modelos empiezan a tener mejoras, con un mayor nivel de confianza, 120 días antes, período de actualización para el que comienzan a incorporarse aquellas variables mensuales con menor rezago que informan sobre el trimestre a proyectar.

Si se analiza la progresión de pronósticos de acuerdo con el calendario de actualización de las variables, se observa que dentro de las proyecciones a uno y dos pasos -salvo para Gets_D- todos los modelos reducen sus RMSFEs conforme se incorpora la nueva información mensual, lo que evidencia un uso eficiente de la información disponible. En el Gráfico 1 se aprecia lo antes mencionado, quedando clara la considerable reducción de los RMSFEs (entre un 50 y 61%) cuando se pasa de predecir de dos pasos a uno.

Gráfico 1. Evolución de los RMSFEs



Nota: Las líneas muestran cómo evolucionan -para cada uno de los modelos- los RMSFE a medida que se aproxima la publicación del dato oficial de Cuentas Nacionales.

Fuente: Elaboración propia

Al dejar de comparar los RMSFEs respecto al *benchmark* y hacerlo entre modelos, Lasso_E, Lasso_D y Gets_E son los que presentan predicciones más acertadas. Si se compara la evolución de los pronósticos, se ve como estos tres modelos y Gets_D -modelos que usan en su especificación rezagos de la variable dependiente- son menos sensibles a la incorporación de los últimos datos (a 60 y 30 días de la publicación), respecto a los modelos de Factores Dinámicos y EP_BR¹⁴. Este resultado puede deberse a que los modelos que usan rezagos del PIB sean los que mejor predicen las variables mensuales a trimestralizar y, por lo tanto, cuando llegan los datos efectivos el ajuste es mejor. O bien, puede deberse a que esos rezagos capturan gran parte de la varianza del modelo, haciendo que las actualizaciones de las variables independientes sean de menor impacto.

Con relación a los antecedentes nacionales, los resultados obtenidos son consistentes en el sentido de que ambos modelos presentan mejor *performance* que los *benchmark*, y el modelo FD de Rodríguez (2014) predice con mayor precisión que el modelo EP directo de Brum y Rodríguez (2016). Sin embargo, dado que no se usan exactamente las mismas variables, los períodos de estimación y evaluación son distintos, y que existen algunas diferencias en las metodologías de estimación, no es posible comparar la precisión de los modelos estimados en este trabajo con la de los antecedentes documentados.

Si en lugar de los RMSFEs, se usa como parámetro de evaluación la frecuencia relativa de veces que cada modelo predice con mínimo error cada uno de los trimestres del período de evaluación, hay consistencia en los resultados, siendo los modelos basados en EP los que presentan mejor desempeño. En la Tabla 4 se aprecia que el modelo Gets_E es el que más veces tiene el menor error para las proyecciones 30, 60 y 180 días antes del dato oficial; Gets_D en las que se hacen 90 días antes, mientras que Lasso_D lo es para 120 y 150 días antes.

¹⁴ Esto se observa tanto en la Tabla 3 como en las pendientes de las líneas del Gráfico 1.

Tabla 4. Frecuencia relativa de predicción con error mínimo, en porcentajes

Modelo	Horizonte de predicción					
	Dos pasos (t+2)			Un paso (t+1)		
	Días antes de la publicación de los datos oficiales					
	180	150	120	90	60	30
Gets_E	21,7	23,5	20,7	20,7	29,6	20,9
Gets_D	16,9	16,5	8,0	23,2	14,8	11,6
Lasso_E	12,0	9,4	14,9	7,3	7,4	9,3
Lasso_D	20,5	24,7	27,6	14,6	6,2	17,4
DF_A	9,6	10,6	10,3	11,0	14,8	12,8
EP_BR	12,0	9,4	9,2	13,4	17,3	15,1
FD_R	7,2	5,9	9,2	9,8	9,9	12,8

Nota: Los valores expresan la frecuencia relativa en la que cada modelo predice con menor error absoluto para cada uno de los trimestres que se predicen, agrupados por el período de actualización.

Fuente: Elaboración propia.

Los resultados presentados en esta sección, que sugieren que los modelos puente superan a los de Factores Dinámicos, difieren a los habituales en la literatura del *Nowcasting*. Lo más frecuentemente hallado en la literatura revisada es que la relación sea inversa, a tal punto de que en numerosos trabajos se utiliza como *benchmark* modelos EP (Angelini et al., 2011; Bańbura et al., 2013; Kunovac y Špalat, 2014; Modugno et al., 2016; Chernis y Sekkel, 2017; Andreini et al., 2020). Los motivos del hallazgo podrían estar asociados a tres factores: a las decisiones que se tomaron respecto a utilizar *autoarima* para prolongar las variables explicativas mensuales, a que la selección de variables sea recursiva y/o a la eficiencia de los métodos GETS y LASSO.

IV.2 Contribución de cada variable a la actualización de los pronósticos

A partir de esta sección, se consideran tres modelos, uno de cada una de las tres principales metodologías aplicadas en el presente trabajo, seleccionados en base a su habilidad predictiva. Dentro de los modelos basados en EP que usan GETS se selecciona Gets_E. En los EP que utilizan LASSO, se trabaja con Lasso_D, modelo que si bien en promedio arroja RMSFEs algo superiores a los de Lasso_E, demuestra

ser el que mejor incorpora los primeros datos mensuales del trimestre a predecir. Del conjunto de los modelos de FD se utiliza FD_A.

Cuando se realizan, tal como en este caso, ejercicios de proyección en *pseudo* tiempo real usando modelos FD y basados en EP, es posible calcular las contribuciones de cada una de las variables a la actualización de los pronósticos.

Para esto, una vez que se proyectan las variables mensuales no disponibles para la totalidad del horizonte de predicción (mediante el filtro de Kalman en los FD y mediante modelos univariados en los basados en EP), se comparan los pronósticos del PIB trimestral que surgen con variables mensuales proyectadas, respecto a las basadas en los verdaderos datos observados.

Siguiendo a Cascaldi-Garcia et al. (2021), se calcula el promedio de la contribución a la predicción de la tasa de crecimiento del PIB trimestral de la variable i (CVP), en el momento de actualización de pronóstico r como :

$$CVP_r^i = \frac{\sum_{t=1}^T \left| E \left[\hat{y}_t^q | X_{v_{r(i)}} \right] - E \left[\hat{y}_t^q | X_{v_{r+30}} \right] \right|}{T}, \quad (9)$$

donde T es el número de observaciones, y_t^q es la tasa de crecimiento interanual del PIB trimestral para t , variable a predecir, r hace referencia al período de actualización de los pronósticos, que en este caso son 30, 60, 90, 120 y 150 días antes de la publicación del dato oficial; $E \left[\hat{y}_t^q | X_{v_{r+30}} \right]$ es la predicción de y_t^q usando la información disponible en el período de actualización $r + 30$; mientras que $E \left[\hat{y}_t^q | X_{v_{r(i)}} \right]$ es la predicción de y_t^q usando la información de i disponible en el momento r , pero usando para las $N - 1$ variables restantes la disponible en $r + 30$, o sea 30 días antes.

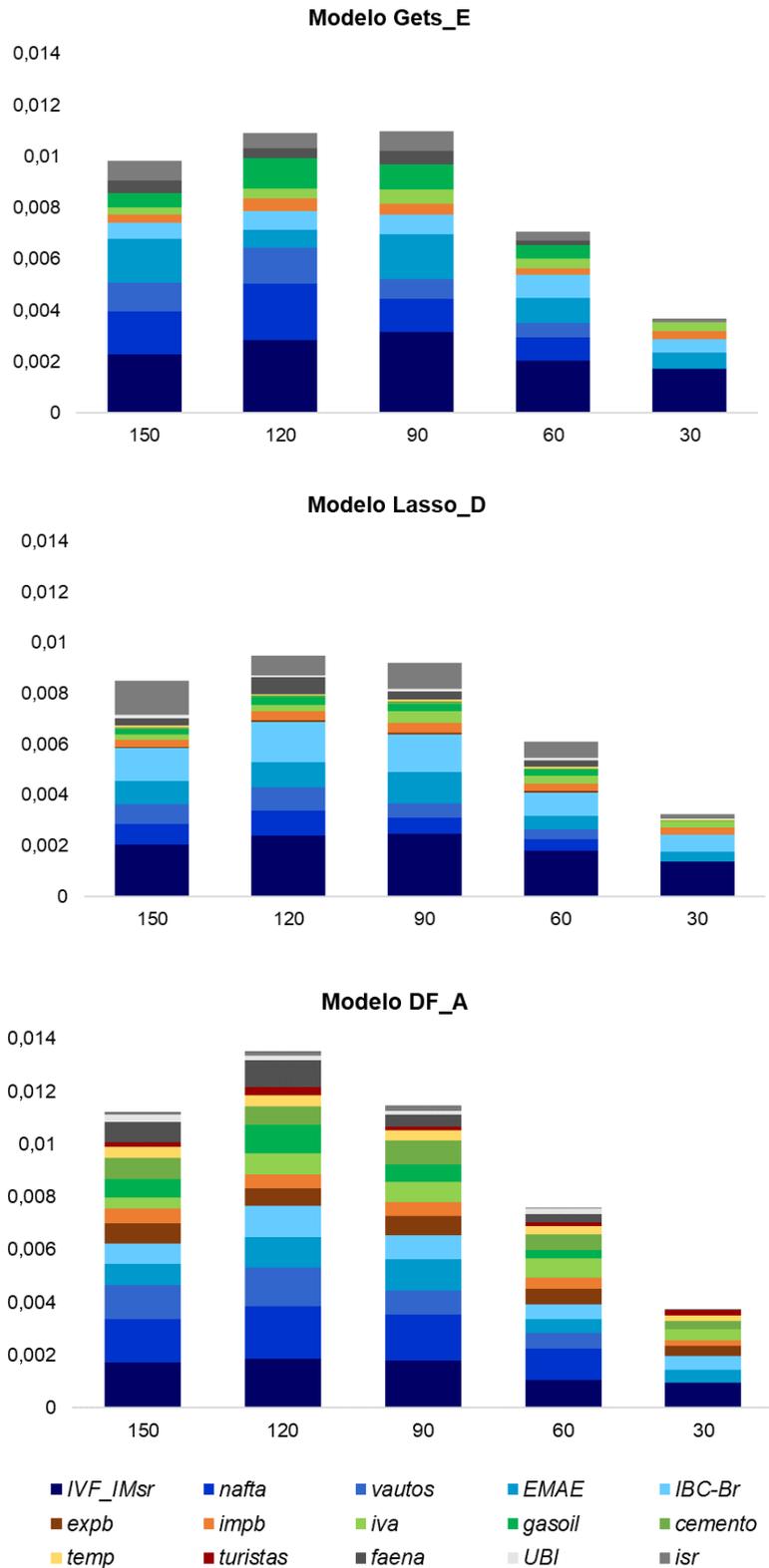
Por ejemplo, la contribución de la tasa de empleo a la actualización de pronósticos que se hace 30 días antes de la publicación del dato oficial (CVP_{30}^{temp}) mide el promedio de la revisión absoluta de la tasa de crecimiento interanual del PIB que se debe a la sustitución de los datos mensuales observados que se encuentran disponibles 30 días

antes, por los que eran proyectados con la información disponible 60 días antes. Nótese que entre $E[\hat{y}_t^q | X_{v_{r(i)}}]$ y $E[\hat{y}_t^q | X_{v_{r+30}}]$ la única variable que difiere es i , teniéndose un dato más de la misma -que deja de ser proyectado- en $E[\hat{y}_t^q | X_{v_{r(i)}}]$ respecto a $E[\hat{y}_t^q | X_{v_{r+30}}]$. Por lo anterior, esta contribución también es conocida como “novedad”, ya que mide el impacto en las proyecciones del PIB asociado a la diferencia entre lo que se espera que sean las variables mensuales y el verdadero dato.

El análisis de estas contribuciones resulta interesante especialmente luego de los resultados presentados en la sección previa, que evidencian que los pronósticos mejoran cuando se incorporan las variables mensuales. En este sentido, este análisis complementa al anterior, ya que identifica a las variables que más inciden en las actualizaciones y por ende, en las mejoras del pronóstico.

En el Gráfico 2 se presentan las CVP para cada una de las variables, agrupadas por el período de actualización, para los tres modelos preseleccionados en la sección anterior.

Gráfico 2. Contribución de las variables a la actualización de los pronósticos



Nota: para cada período de actualización de pronóstico, se presenta la contribución de cada una de las variables en términos absolutos a la revisión de los pronósticos. El eje horizontal muestra los períodos de actualización de pronósticos. El eje vertical presenta los valores absolutos de las actualizaciones de los pronósticos sobre las tasas de variación interanuales.

Fuente: Elaboración propia.

Una primera apreciación es que el modelo FD_A presenta para cada uno de los períodos de actualización, valores para las 15 variables, mientras que para los restantes modelos, hay variables que no registran valores (no se muestran en las barras). Esto responde a que en los FD, todas las variables forman parte del factor en toda la muestra, mientras que en los basados en EP, las variables son seleccionadas recursivamente, habiendo algunas que se seleccionan muchas veces, otras pocas y otras ninguna.

A pesar de lo anterior, y de las diferencias en las metodologías de estimación, se puede observar cómo en los tres modelos hay cinco variables que contribuyen a explicar -para todos los períodos de actualización- más del 50%¹⁵ de las revisiones de los pronósticos. Estas variables, que aparecen en tonos de azul, son: el Índice de Volumen de la Industria Manufacturera sin refinería, las ventas de autos, las ventas de nafta, y las variables externas que refieren a las economías de Argentina y Brasil. Esto permite inferir que a lo largo del período de evaluación estas son las variables que más contribuyen a la mejora de los pronósticos.

Este resultado podría deberse a particularidades de la relación de dichas variables con el PIB. Específicamente, *IVF_IMsr* además de ser una buena *proxy* del valor agregado de la Industria Manufacturera (cuyo peso en el PIB en 2016 era de 11%), constituye un buen indicador de exportaciones, ya que incluye la actividad de las principales empresas exportadoras de bienes del país (las que operan en Zonas Francas y los frigoríficos). A su vez, al ser la industria una actividad que demanda insumos, tanto nacionales (principalmente agropecuarios) como importados, podría también aproximar la dinámica del sector agropecuario, así como de las importaciones. Por su parte, *vautos* puede ser relevante, además de por su relación con el valor agregado del sector comercio, por su vinculación con la percepción general que tiene la población respecto al nivel de actividad, ya que para adquirir estos

¹⁵ En los modelos Gets_E y Lasso_S, los valores superan para todos los períodos de actualización el 70%, siendo los valores más bajos, los que se obtienen el FD_A.

bienes es necesario decidir endeudarse a mediano plazo o usar ahorros, lo que es probable que esté influido por dicha percepción. En lo que respecta a *nafta*, se trata de una variable asociada con la movilidad, por lo que constituye un buen indicador de aquellas actividades que requieren presencialidad. Por último, la importancia de *EMAE* e *IBC-Br* se vincularía a la relevancia que tienen Argentina y Brasil en el período de análisis, como demandantes de bienes uruguayos, así como también de servicios, entre ellos los asociados al turismo.

Otro aspecto destacable a partir del cálculo de las CVP, y que se observa gráficamente a partir del largo de las barras, es que las actualizaciones a causa de los datos mensuales que arriban 150, 120 y 90 días antes de la publicación, son en promedio de mayor magnitud que las ocurridas 60 días antes y más aún que 30 días antes de la publicación. Esto podría deberse, además del hecho de que cada vez son menos los meses a proyectar para completar el horizonte de predicción, a que las proyecciones de las variables mensuales sean más erráticas para horizontes de predicción mayores, por lo que a medida que se sustituyen datos por pronósticos, las revisiones de las variables que informan sobre el trimestre en curso se reducen, resultando en menores actualizaciones.

Por último, si en lugar de cotejar entre períodos, se compara entre los modelos, se observa que las barras de *FD_A* son mayores para los cinco períodos de actualización. Esto va en línea con los resultados de la sección anterior, donde se muestra que los modelos que no usan rezagos del PIB trimestral presentan mejoras más pronunciadas en el desempeño predictivo a medida que se incorporan los nuevos datos mensuales.

IV.3 Pronósticos en momentos de grandes quiebres

El análisis del desempeño predictivo de los modelos *de Nowcasting* en circunstancias en que la actividad sufre importantes cambios puede ser relevante por varios motivos. En estos momentos, las decisiones que se toman tanto por parte de agentes privados

como, principalmente, por los hacedores de política económica, resultan relevantes ya sea para la minimización de los efectos adversos, como para la reducción de costos y plazos en la recuperación. Debido a que para la toma de estas decisiones es fundamental tener información oportuna y certera, cuando los modelos de *Nowcasting* brindan estimaciones precisas, resultan herramientas útiles para conocer el estado de situación de la economía aún en estos momentos, e intentar tomar acciones correctivas o mitigatorias.

En los últimos veinte años la economía uruguaya ha atravesado dos grandes episodios en los que la actividad económica ha sufrido caídas abruptas: la crisis del 2002 y la del 2020. Si bien los efectos de esas crisis no tuvieron lugar en un único trimestre en esos años, es posible identificar al tercer trimestre de 2002 y al segundo de 2020 como los más afectados¹⁶, por lo que esta sección se encuentra enfocada en el análisis de esos períodos.

IV.3.1 Crisis del 2002

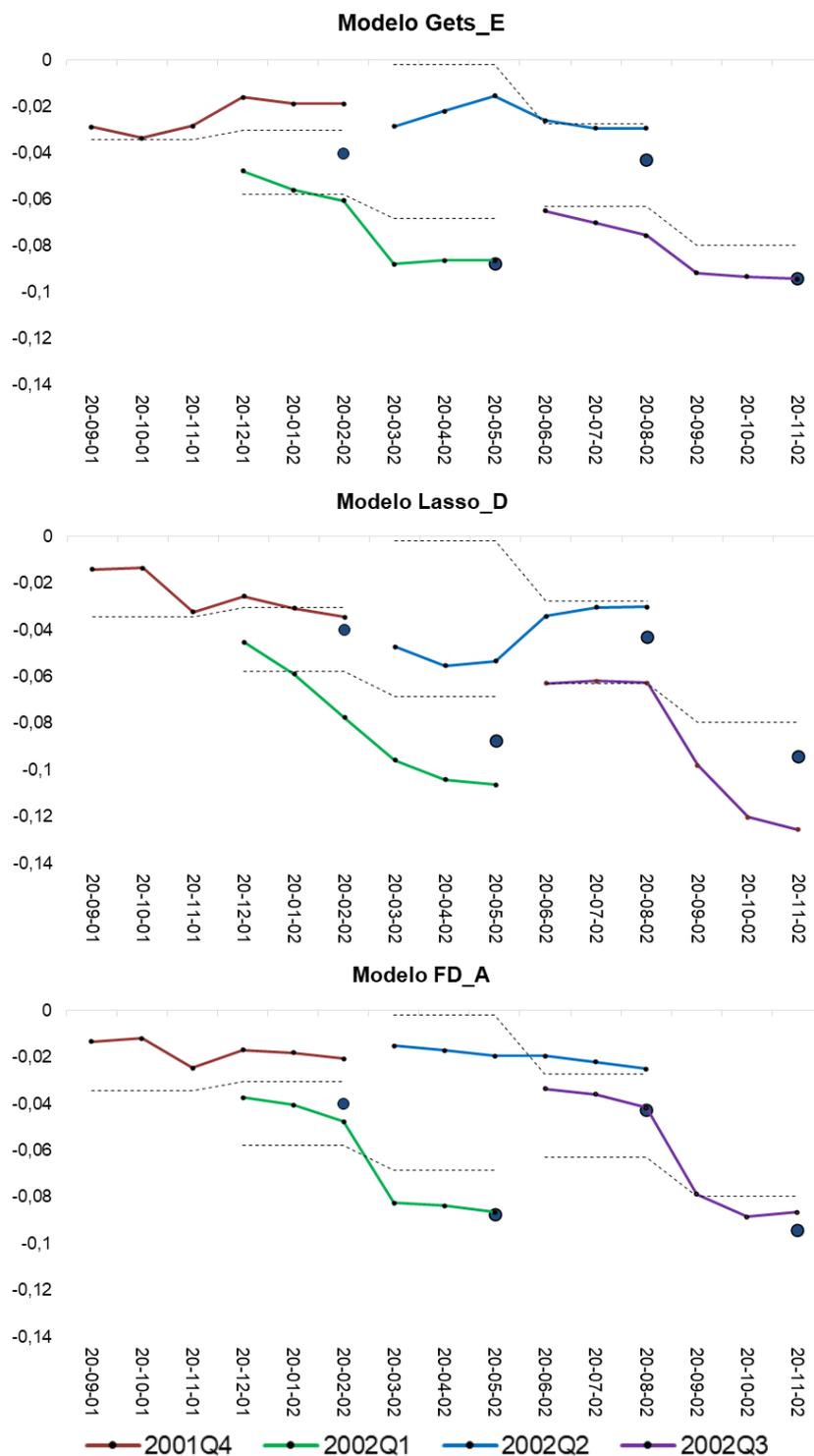
A inicios del año 2002 la economía uruguaya ya presenta signos de fragilidad, habiendo registrado en los tres años previos a 2002 caídas en el nivel de actividad. Si bien en esta situación inciden diversos factores, internacionales y locales, sintéticamente, las causas pueden resumirse en los siguientes hechos: a inicios de 1999, Brasil -principal socio comercial junto con Argentina- devalúa su moneda, generando una importante pérdida de competitividad bilateral. Posteriormente, a finales de 2001 Argentina atraviesa una crisis financiera con un posterior estallido social, que repercute negativamente en Uruguay, principalmente en los flujos de comercio bilateral con dicho país y en el sistema bancario, por el importante volumen de depósitos de argentinos en el sistema bancario uruguayo. En ese contexto, al que

¹⁶ 2002Q3 y 2020Q2 son los trimestres en los que la actividad económica presenta mayores caídas interanuales para la serie trimestral del PIB que abarca 1994-2021. A su vez, haciendo una detección de quiebres para dicha serie con la función Impulso Saturación de GETS con un 99% de confianza, ambos trimestres son los que resultan significativos como atípicos.

se la suma la detección de fiebre de aftosa en 2001, una caída en los precios de *commodities* exportados por Uruguay y una débil situación macroeconómica, es que en julio de 2002, luego de meses de masivos retiros en los depósitos bancarios, se produce con la declaración de un feriado bancario el estallido de la crisis (Fernández et al., 2003; Bucheli y Furtado 2004; Mordecki, 2017) . Esta crisis con origen financiero rápidamente tuvo su impacto en el sector real, con caídas en el empleo, los salarios, aumento de la pobreza, entre otros efectos. En 2002 se registra una caída del PIB de 7,9% y de 9,4% en el tercer trimestre del año.

En el Gráfico 3 se presentan, para cada uno de los modelos, las sucesivas proyecciones que surgen del ejercicio en *pseudo* tiempo real del tercer trimestre de 2002 y los tres trimestres que lo anteceden.

Gráfico 3. Actualización de pronósticos, crisis del 2002



Nota: En los tres modelos, las líneas continuas muestran las seis predicciones que se realizan para cada uno de los trimestres del período 2001Q4-2002Q3. Las líneas punteadas indican las predicciones que realiza el mejor de los modelos benchmark (autoarima). Los puntos azules indican los datos observados del PIB trimestral. Las fechas del eje horizontal refieren al momento en el que se realiza cada uno de los pronósticos. Todos los valores están expresados en tasas de crecimiento interanual.

Fuente: Elaboración propia.

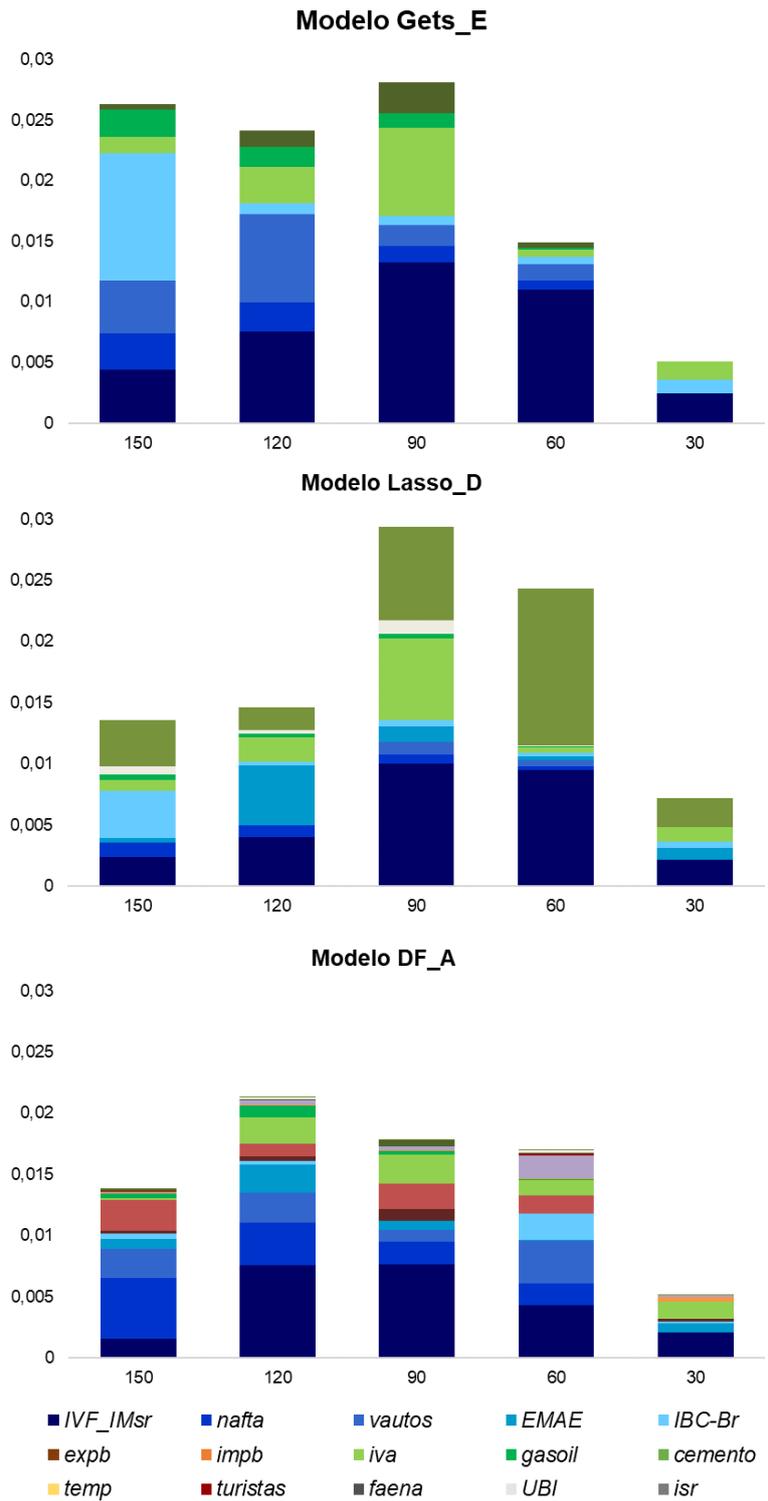
Al comparar la *performance* predictiva de los tres modelos para 2002Q3 respecto al *benchmark*, los resultados son dispares. Gets_E es el único de los tres cuyas predicciones superan la capacidad del *benchmark*, tanto para uno como para dos pasos. Lasso_D, por su parte, presenta errores similares al *benchmark* cuando se proyecta a dos pasos, pero cuando se conoce el dato de 2002Q2 y se pasa a estimar a uno, a pesar de que en la estimación a 90 días el error es pequeño, en las predicciones 60 y 30 días antes se sobreestima la caída del nivel de actividad. Tanto es así que los errores superan en términos absolutos los del *benchmark*. En tanto, el modelo FD_A muestra mejores predicciones que el *benchmark* solo en las predicciones 60 y 30 días antes.

Si se analiza cómo repercute en los pronósticos de 2002Q3 la incorporación de nuevos datos mensuales, Gets_E parece ser el que hace un mejor uso de los mismos. Este modelo mejora notoriamente las predicciones cuando predice 120 y 150 días antes respecto a cuando lo hace 150 y 180, mientras que en las estimaciones a un paso, la llegada de nuevos datos no genera grandes revisiones de la proyección a 90 días (que resulta ajustada). Las predicciones de Lasso_D conforme se incorporan los nuevos datos mensuales son muy poco sensibles para $t+2$, opuesto a lo que sucede cuando se predice para $t+1$. Por su parte, el FD_A se muestra poco sensible al incorporar nuevos datos mensuales, pero la incorporación de información siempre mejora los pronósticos, salvo cuando se pasa de predecir de 60 días antes, a 30.

De forma análoga a la sección IV.2, resulta de interés analizar cuál es la contribución de cada una de las variables a la actualización de los pronósticos, no solo para el promedio del periodo de evaluación, sino también para la sucesión de predicciones referidas al 2002Q3.

En el Gráfico 4, al igual que para el promedio -presentado en el Gráfico 2- predominan los tonos de azul, de modo que las cinco variables que más inciden en la media del período son también las fundamentales a la hora de explicar las revisiones de los pronósticos en este trimestre de fuerte caída.

Gráfico 4. Contribución de las variables a la actualización de los pronósticos, crisis del 2002



Nota: para cada período de actualización de pronóstico, se presenta la contribución de cada una de las variables en términos absolutos a la actualización de los pronósticos del 2002Q3. El eje horizontal muestra los períodos de actualización de pronósticos. El eje vertical presenta los valores absolutos de las actualizaciones de los pronósticos sobre las tasas de variación interanuales.

Fuente: Elaboración propia.

Si se observa la composición dentro de estas cinco variables, destaca en los tres modelos el alto peso de *IVF_IMsr*, mayor que para el resto del período. Por su parte, las incidencias de las variables *EMAE* Y *IBC-Br*, varían entre modelos y períodos, no pudiéndose encontrar un patrón, sobresaliendo la alta incidencia de *IBC-Br* para las actualizaciones 150 días antes en *Gets_E* y *Lasso_D*, y de *EMAE* para *Lasso_D* 120 días antes de la publicación del dato oficial.

Respecto a las variables que se encuentran por fuera de esas cinco, *iva* presenta un notorio aumento de ponderación en los tres modelos cuando se actualizan los pronósticos 90 días antes, que dado su rezago de 30 días, se debe a la incorporación de los datos del mes de julio 2002. Por su parte, en *Lasso_D* el peso de la contribución de *isr* resulta relativamente alta en todos los momentos de actualización. Esta incidencia, teniendo en cuenta que ese indicador sufrió en 2002 una caída cercana al 20%, parece explicar por qué los pronósticos sobreestiman la caída del PIB.

En síntesis, se aprecia que el modelo *Gets_E* es en términos relativos el mejor al predecir 2002Q3. Las predicciones de este modelo resultan muy ajustadas, en tanto que al 20 de julio de 2002 (150 días antes de la publicación del PIB) el modelo predice una caída interanual para 2002Q3 de 7,02%, con un error de predicción asociado de 2,4 puntos porcentuales, que disminuye en las subsiguientes proyecciones, prediciendo al 20 de noviembre casi exactamente el dato observado.

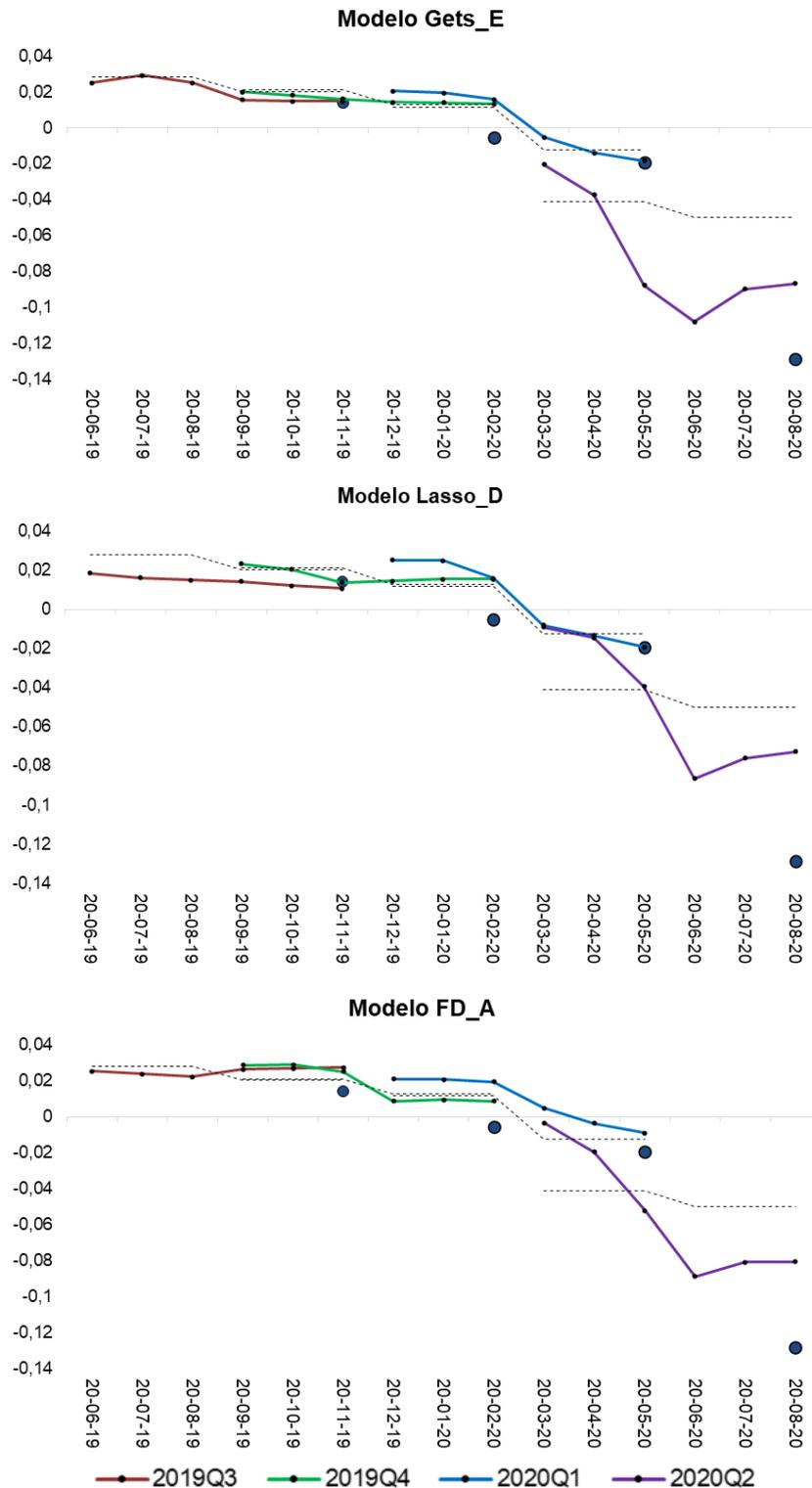
IV.3.2 Crisis de 2020

En 2020, la economía uruguaya atraviesa la recesión de carácter global generada por el Covid-19. Al igual que en muchos otros países, se registra una abrupta caída del nivel de actividad consecuencia de las medidas adoptadas (tanto interna como externamente) en orden de prevenir los contagios. Entre las medidas adoptadas en Uruguay, cabe mencionar la suspensión de las clases presenciales para todos los niveles de enseñanza, el incentivo al teletrabajo, el cierre de fronteras y la suspensión de la actividad de restaurantes, centros deportivos, centros comerciales, culturales y

artísticos (Capurro et al., 2020). La mayoría de estas medidas empiezan a regir desde la segunda quincena de marzo. Tales decisiones, en interacción con las tomadas en el resto de los países, contribuyen a que en 2020 la caída del PIB uruguayo sea de 5,8%, y de un 12,8% en el segundo trimestre del año.

En el Gráfico 5 se presenta para cada uno de los modelos las sucesivas proyecciones que surgen del ejercicio en *pseudo* tiempo real del segundo trimestre de 2020 y los tres trimestres que lo anteceden.

Gráfico 5. Actualización de pronósticos, crisis de 2020



Nota: En los tres modelos, las líneas continuas muestran las seis predicciones que se realizan para cada uno de los trimestres del período 2019Q3-2020Q2. Las líneas punteadas indican las predicciones que realiza el mejor de los modelos benchmark (autoarima). Los puntos azules indican los datos observados del PIB trimestral. Las fechas del eje horizontal refieren al momento en el que se realiza cada uno de los pronósticos. Todos los valores están expresados en tasas de crecimiento interanual.

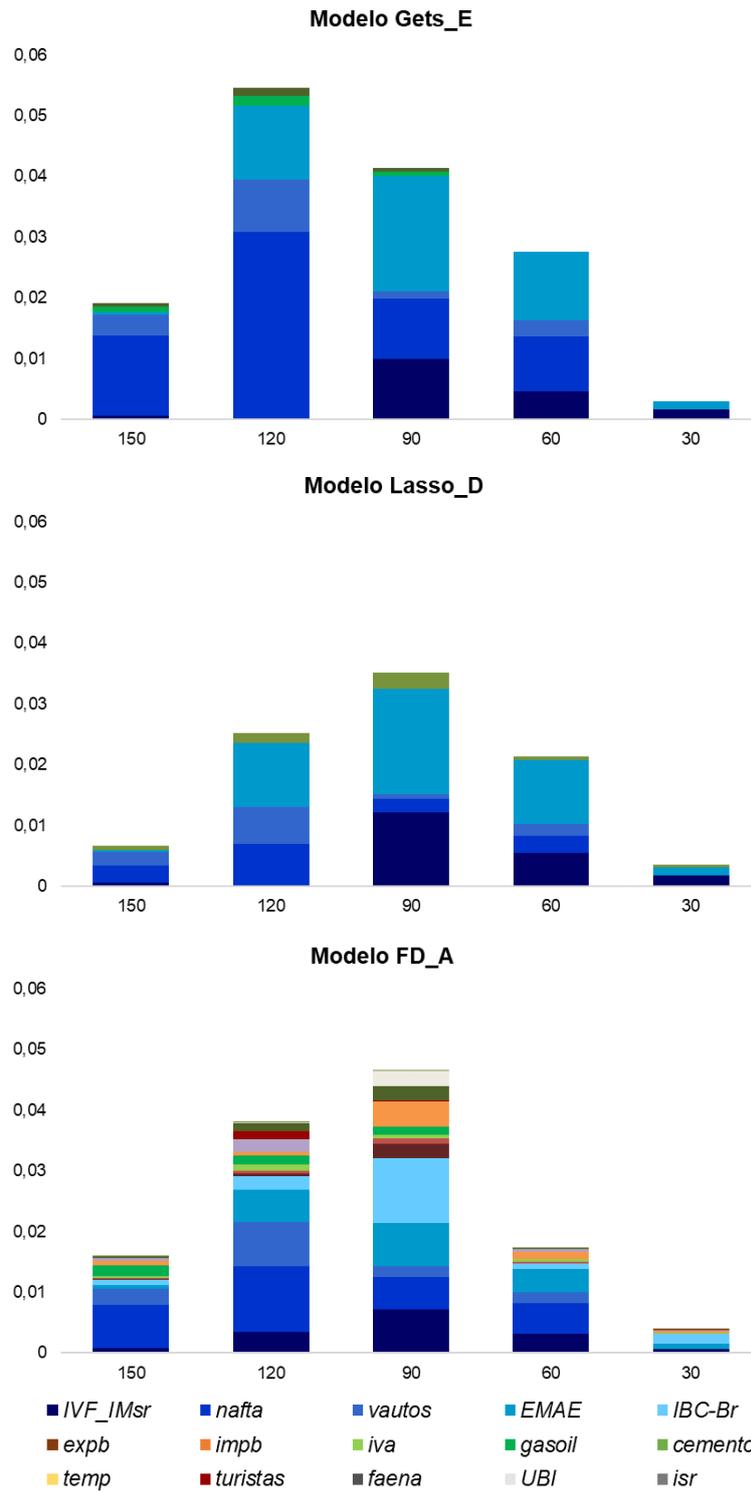
Fuente: Elaboración propia.

Si se compara la capacidad predictiva de los modelos para 2020Q2 en relación al *benchmark*, se encuentra que los tres modelos muestran una mejor *performance* desde el pronóstico realizado 150 días antes a la publicación, siendo el Gets_E el modelo con menores errores de predicción para los seis períodos de actualización.

Al analizar como incide la incorporación de nuevos datos mensuales en los modelos, hay un claro patrón que se repite en los tres: disminución de los errores conforme se incorporan las nuevas variables hasta 90 días antes de la publicación, seguida de un aumento de los mismos en las estimaciones 60 y 30 días antes de la publicación del dato referente a 2020Q2.

En el Gráfico 6 se considera la contribución de cada una de las variables en la actualización de los pronósticos para 2020Q2. Se aprecia que nuevamente predominan los tonos de azul, correspondientes a las cinco variables ya destacadas. Esto demuestra que dichas variables son las que más inciden en la actualización de los pronósticos de este trimestre, al igual de lo que sucedía en el tercer trimestre de 2002 y en el promedio del período estudiado.

Gráfico 6. Contribución de las variables a la actualización de los pronósticos, crisis del 2020



Nota: para cada período de actualización de pronóstico, se presenta la contribución de cada una de las variables en términos absolutos a la actualización de los pronósticos del 2020Q2. El eje horizontal muestra los períodos de actualización de pronósticos. El eje vertical presenta los valores absolutos de las actualizaciones de los pronósticos.

Fuente: Elaboración propia.

A diferencia de lo ocurrido para 2002Q3 el indicador de *IVF_IMsr* no tiene gran incidencia en las actualizaciones, salvo en la actualización 90 días antes.

En el modelo *Gets_E* hay dos variables que son las que más inciden en la actualización de la secuencia de pronósticos. La primera es *nafta*, cuyo principal aporte se encuentra en los pronósticos 150 y 120 días antes, períodos donde dado el rezago de esta variable se incorporan los datos de los meses marzo y abril, que es cuando empiezan a regir las medidas asociadas al Covid-19. La incidencia de esta variable, que sigue siendo relevante aunque en menor medida en las sucesivas actualizaciones, captura principalmente la reducción de la movilidad asociada a las medidas tomadas. En segundo lugar, es elevada la ponderación del *EMAE*, principalmente en las actualizaciones 120, 90 y 60 días antes. En dichos períodos se incorporan los datos de la actividad para Argentina correspondiente a los meses de marzo, abril y mayo. Por su parte, la variable *vautos*, tiene una contribución alta, aunque bastante menor a *nafta*, en las actualizaciones 150 y 120 días antes, donde se incorporan sus datos que informan sobre los meses de marzo y abril.

La principal diferencia entre el *Gets_E* y *Lasso_D* reside en la variable *nafta*, que contribuye en menor medida en el segundo modelo. En tanto, las variables *EMAE* y *vautos* presentan un comportamiento similar. En este modelo, al igual que para 2002Q3, *isr* presenta una incidencia relevante, siendo la única variable preseleccionada para predecir este trimestre en los dos modelos basados en EP, por fuera de las cinco que más inciden.

Por último, en el modelo *FD_A* las contribuciones a la actualización de pronósticos de las variables *nafta*, *EMAE* y *vautos*, son similares a la del *Gets_E*. La principal diferencia reside en la variable *IBC-Br*, que no es seleccionada por los modelos basados en EP para predecir 2020Q2 pero muestra gran contribución en la actualización 90 días del *FD_A*.

En este caso, a diferencia de 2002Q3, los tres modelos presentan un patrón de actualización similar, siendo Gets_E nuevamente el que tiene un mejor desempeño. Esto último se debe, en particular, a la mejor incorporación de información mensual en las predicciones a dos pasos.

V. SÍNTESIS DE RESULTADOS

El primer objetivo del trabajo consistió en estimar modelos de *Nowcasting* que brinden información sobre la situación económica a tiempo. Para ello, a partir de un conjunto de indicadores con frecuencia mensual se estimaron modelos para predecir el PIB trimestral de Uruguay. En particular, se plantearon dos tipos de modelos: modelos FD y basados en EP, que fueron evaluados mediante un ejercicio en *pseudo* tiempo real; realizando para cada trimestre seis pronósticos, el primero de ellos 180 días antes de la publicación del dato de Cuentas Nacionales.

En términos generales, los resultados obtenidos se encuentran en línea con la literatura. En primer lugar, los modelos propuestos cuentan con un mejor desempeño relativo respecto a modelos *benchmark* tomados como referencia. En segundo, la gran mayoría de los modelos estimados mejoran las proyecciones conforme incorporan la nueva información mensual, denotando un uso eficiente de la misma.

Al comparar la capacidad predictiva entre los modelos, se obtuvo como resultado que modelos basados en EP tienen mejor desempeño predictivo que los FD. Este hallazgo, que no es habitual en la literatura, podría deberse a que en el presente trabajo se estimaron EP en los que las variables fueron seleccionadas de forma recursiva a lo largo del período de análisis, mediante los métodos LASSO y GETS. La selección recursiva de las variables permite que los modelos sean más flexibles y que puedan adaptarse a los cambios en las dinámicas relativas entre las variables.

El segundo objetivo, consistió en analizar para tres modelos seleccionados en base a su capacidad predictiva, la contribución de cada una de las variables en la actualización de los pronósticos. Como resultado, se encuentra que para los tres modelos y en los seis períodos de actualización, hay cinco variables que explican más del 50% de las revisiones absolutas de los pronósticos: El Índice de Volumen Físico de la Industria Manufacturera sin refinería, la venta de automóviles 0 km, la venta de nafta, y las variables que informan sobre la actividad de Argentina y Brasil.

El tercer y último objetivo consistió en evaluar la capacidad predictiva en la crisis del 2002 y la asociada al Covid-19. Si bien se trata de quiebres alejados en el tiempo y que responden a causas muy distintas, el modelo basado en EP en el que las variables se seleccionaron mediante GETS y usa una muestra con ventana expandible, es el que mejor desempeño presenta en ambos momentos. Esto podría estar asociado con la flexibilidad para la selección de las variables y con una mayor eficiencia del GETS frente al LASSO para incorporar aquellas que mejor representen la dinámica de la actividad económica en momentos de inestabilidad.

VI. CONSIDERACIONES FINALES

El presente trabajo implica una contribución a la literatura empírica nacional por varios aspectos. En primer lugar, se estiman un conjunto de modelos de *Nowcasting* que presentan un buen desempeño para obtener estimaciones oportunas que anticipen el crecimiento del PIB trimestral. Los modelos estimados pueden considerarse actualizaciones de los modelos documentados en los antecedentes nacionales, ya que se les incorporan las nuevas series de Cuentas Nacionales base 2016, nuevos indicadores mensuales, e innovaciones respecto a la forma de seleccionar y proyectar las variables en los modelos basados en Ecuaciones Puente.

En segundo lugar, el análisis de la contribución de las variables a la actualización de los pronósticos resulta una novedad en la literatura nacional, que permite identificar cuáles son los indicadores con frecuencia mayor a la trimestral que mejor aproximan el comportamiento del PIB, tanto en momentos ordinarios, como en los momentos en los que la economía atraviesa fluctuaciones abruptas.

En tercer lugar, el estudio del desempeño de los modelos en las crisis de 2002 y 2020 permite identificar algunas ventajas y limitaciones que tiene el uso de estos modelos en momentos en que la actividad económica sufre cambios abruptos. Si bien los modelos planteados presentan mejoras respecto a los *benchmark*, los errores de predicción en estos momentos son de mayor magnitud que en momentos ordinarios, por lo que es necesario más investigación sobre el comportamiento de los modelos de *Nowcasting* en estos casos.

Desde el punto de vista metodológico, la principal contribución respecto a la literatura revisada consistió en la propuesta de recursividad para la selección de las variables empleadas en los modelos EP.

Para finalizar, a partir de este trabajo, surgen distintas líneas de investigación posibles. Por un lado, se podrían incluir a los ejercicios de estimación y evaluación, modelos que utilizan las metodologías MIDAS y Vectores Autorregresivos en Frecuencia Mixta,

para explorar si existen diferencias relevantes con los resultados obtenidos para los modelos FD y EP. Por otro lado, para períodos más recientes (que no abarquen la crisis de 2002) se podrían plantear modelos de *Nowcasting* con un mayor número de variables, explotando la cada vez mayor cantidad de información disponible.

BIBLIOGRAFÍA

- Alvarez, R., Camacho, M. y Pérez-Quirós, G. (2012). *Finite Sample Performance of Small Versus Large Scale Dynamic Factor Models*. CEPR Discussion Paper. 8867.
- Álvez, M., Bucacos, E., Mateauda, M. y Pienika, E. (2021). *Retropolación para series de Cuentas Nacionales Trimestrales. Series de Producto Interno Bruto de Uruguay con frecuencia trimestral para el período 1997-2015*. Documentos de Trabajo. Banco Central del Uruguay.
- Andreini, P., Senftleben-König, C., Hasenzagl, T., Reichlin, L. y Strohsal, T. (2020). *Nowcasting German GDP*. SSRN Scholarly Paper. ID 3526048. Rochester, NY: Social Science Research Network.
- Angelini, E., Camba-Mendez, G., Giannone, D., Reichlin, L., y Rünstler, G. (2011). *Short-term forecasts of euro area GDP growth*. *The Econometrics Journal* 14(1):25-44.
- Baffigi, A., Golinelli, R. y Parigi, G. (2004). *Bridge models to forecast the euro area GDP*. *International Journal of Forecasting* 20(3):447-60.
- Bai, J., y Ng, S. (2002). *Determining the Number of Factors in Approximate Factor Models*. *Econometrica* 70(1):191-221. doi: 10.1111/1468-0262.00273.
- Bañbura, M., Giannone, M., Modugno, M. y Reichlin, L. (2013). *Now-casting and the real-time data flow*. Pp. 195-237 en *Handbook of economic forecasting*. Vol. 2. Elsevier.
- Bañbura, M., Giannone, M., Modugno, M. y Reichlin, L. (2010). *Nowcasting*. ECB Working Paper No. 1275.
- Bañbura, M. y Modugno, M. (2014). *Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data*. *Journal of Applied Econometrics* 29(1):133-60. doi: 10.1002/jae.2306.

- Belloni, A. y Chernozhukov, V. (2011). *High Dimensional Sparse Econometric Models: An Introduction*. MIT Department of Economics Working Paper No. 11-17.
- Brum, C., y Rodríguez, H. (2016). *Modelos Puente Para Proyectar El PIB En El Corto Plazo. Enfoque Sectorial*. Documentos de Trabajo. Banco Central del Uruguay.
- Bucheli, M., y Furtado, M. (2004). *Uruguay 1998-2002: ¿quiénes ganaron y quiénes perdieron en la crisis?* CEPAL.
- Campos, J., Ericsson, N. y Hendry, D. (2005). *General-to-specific modeling: an overview and selected bibliography*. International Finance Discussion Paper. 838. Board of Governors of the Federal Reserve System (U.S.).
- Capurro, A., Deagosto, G., Ferro, F., Ithurralde, S. y Oddone, G. (2020). *Impacto Social y Económico de La COVID-19 y Opciones de Políticas En Uruguay*. https://www.latinamerica.undp.org/content/rblac/es/home/library/crisis_prevention_and_recovery/impacto-social-y-economico-de-la-covid-19-y-opciones-de-politica.html.
- Cascaldi-Garcia, D., Ferreira, T., Giannone, D. y Modugno, M. (2021). *Back to the Present: Learning About the Euro Area Through a Now-Casting Model*. SSRN Scholarly Paper. ID 3865429. Rochester, NY: Social Science Research Network. doi: 10.17016/IFDP.2021.1313.
- Chernis, T., y Sekkel, R. (2017). *A Dynamic Factor Model for Nowcasting Canadian GDP Growth*. Empirical Economics 53(1):217-34. doi: 10.1007/s00181-017-1254-1.
- Dempster, A., Laird, N. y Rubin, D. (1977). *Maximum Likelihood from Incomplete Data via the EM Algorithm*. Journal of the Royal Statistical Society. Series B (Methodological) 39(1):1-38.
- Diebold, F., y Mariano, R. (1995). *Comparing Predictive Accuracy*. Journal of Business & Economic Statistics 13(3):253-63.

- Dianessi, M., Lanzilotta, B., Mordecki, G., Presa, M., Rodriguez-Collazo, S. (2020). *El crecimiento reciente de Uruguay y desacople de la región: ¿adiós a la semisuma?* Serie Documentos de Trabajo, DT 01/2020. Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.
- Diron, M. (2008). *Short-Term Forecasts of Euro Area Real GDP Growth: An Assessment of Real-Time Performance Based on Vintage Data*. Journal of Forecasting 27(5):371-90.
- Doornik, J. (2009). Autometrics. In JL Castle, N Shephard (eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*:88–121. Oxford University Press, Oxford.
- Doornik, J. y Hendry D. (2007). *Empirical Econometric Modelling – PcGive 12: Volume I*. Timberlake Consultants Ltd., London.
- Doz, C., Giannone, D. y Reichlin, L. (2011). *A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering*. Journal of Econometrics 164(1):188-205. doi: 10.1016/j.jeconom.2011.02.012.
- Doz, C., Giannone, D. y Reichlin, L. (2012). *A Quasi-Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models*. The Review of Economics and Statistics 94(4):1014-24. doi: 10.1162/REST_a_00225.
- Fernández, A., Garda, P. y Perelmuter, P. (2003). *Calificaciones, crisis de deuda y crisis bancaria: la experiencia uruguaya en el 2002*. CINVE.
- Froni, C, y Marcellino, M. (2013). *A Survey of Econometric Methods for Mixed-Frequency Data*. SSRN Scholarly Paper. ID 2268912. Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.2268912.
- Giannone, D. y Reichlin, L. y Sala, L. (2005). *Monetary Policy in Real Time*. NBER Macroeconomics Annual 19. doi: 10.2139/ssrn.680521.

- Giannone, D. y Reichlin, L. y Small, D. (2008). *Nowcasting: The Real-Time Informational Content of Macroeconomic Data*. *Journal of Monetary Economics* 55(4):665-76. doi: 10.1016/j.jmoneco.2008.05.010.
- Hahn, E., y Skudelny, F. (2008). *Early estimates of Euro Area real GDP growth: a bottom-up approach from the production side*. Working Paper Series. 975. European Central Bank.
- Hendry, D. 2006. *Robustifying forecasts from equilibrium-correction models*. *Journal of Econometrics* 135:399-426. doi: 10.1016/j.jeconom.2005.07.029.
- Hendry DF, Doornik J (2014). *Empirical Model Discovery and Theory Evaluation*. The MIT Press, London.
- Hendry, D., Johansen, S. y Santos C. (2007). *Automatic Selection of Indicators in a Fully Saturated Regression*. *Computational Statistics*, 23(2):317–335. doi:10.1007/s00180-007-0054-z.
- Hoover, K. y Perez, S. (1999). *Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search*. *The Econometrics Journal*, 2(2), 167–191. doi:10.1111/1368-423x.00025.
- Hyndman, R. y Khandakar, Y. (2008). *Automatic Time Series Forecasting: The Forecast Package for R*. *Journal of Statistical Software* 27:1-22. doi: 10.18637/jss.v027.i03.
- Ingenito, R. y Trehan, B. (1996). *Using monthly data to predict quarterly output*. *Economic Review* 3:3-11.
- Kunovac, D., y Špalat, B. (2014). *Nowcasting GDP Using Available Monthly Indicators*. The Croatian National Bank, Croatia.
- Kwiatkowski, D, Phillips, P., Schmidt, P. y Shin, Y. (1992). *Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root*. *Journal of Econometrics* 54(1-3):159-78. doi: 10.1016/0304-4076(92)90104-Y.

- Lanzilotta, B., Llambí, C., y Mordecki, G. (2003). *La influencia regional sobre la economía uruguaya: un análisis de los últimos veinte años*. Serie Documentos de Trabajo, DT 01/03. Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.
- Mariano, R., y Murasawa, Y. (2003) *A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series*. Journal of Applied Econometrics 18(4):427-43.
- Masoller, A. (2002). *Indicador sintético de la actividad económica*. Banco Central del Uruguay Jornadas Anuales de Economía 2002.
- Modugno, M, Soybilgen, B. y Yazgan, M. (2016). *Nowcasting Turkish GDP and News Decomposition*. Finance and Economics Discussion Series 2016:1-38. doi: 10.17016/FEDS.2016.044.
- Mordecki, G. (2017) *Uruguay en democracia: treinta años de evolución económica (1985-2015)*. Serie Documentos de Trabajo, DT 08/17. Instituto de Economía, Facultad de Ciencias Económicas y Administración, Universidad de la República, Uruguay.
- Reade, J., y Sucarrat, G. (2016). *General-to-Specific (GETS) Modelling And Indicator Saturation With The R Package Gets*. University of Oxford, Department of Economics.
- Tibshirani, T. (1996). *Regression Shrinkage and Selection via the Lasso*. Journal of the Royal Statistical Society. Series B (Methodological) 58(1):267-88.
- Rodríguez, H. (2014). *Un Indicador de La Evolución Del PIB Uruguayo En Tiempo Real*. Banco Central del Uruguay.
- Sargent, T. y Sims, C. (1977). *Business Cycle Modeling without Pretending to Have Too Much a Priori Economic Theory*. 55. Federal Reserve Bank of Minneapolis.

- Sax, C., Eddelbuettel, D. (2018). *Seasonal Adjustment by X-13ARIMA-SEATS in R*. Journal of Statistical Software 87(11). doi: 10.18637/jss.v087.i11.
- Schumacher, C., y Breitung, J. (2008). *Real-Time Forecasting of German GDP Based on a Large Factor Model with Monthly and Quarterly Data*. International Journal of Forecasting 24(3):386-98. doi: 10.1016/j.ijforecast.2008.03.008.
- Shumway, R., y Stoffer, D. (1982). *An Approach to Time Series Smoothing and Forecasting Using the Em Algorithm*. doi: 10.1111/J.1467-9892.1982.TB00349.X.
- Stock, J. y Watson, M. (1998). *Diffusion Indexes*. SSRN Scholarly Paper. ID 226366. Rochester, NY: Social Science Research Network.
- Stock, J. y Watson, M. (2002). *Forecasting Using Principal Components From a Large Number of Predictors*. JASA. Journal of the American Statistical Association 97. doi: 10.1198/016214502388618960.
- Watson, M. y Engle, R. (1983). *Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models*. Journal of Econometrics 23(3):385-400. doi: 10.1016/0304-4076(83)90066-0.

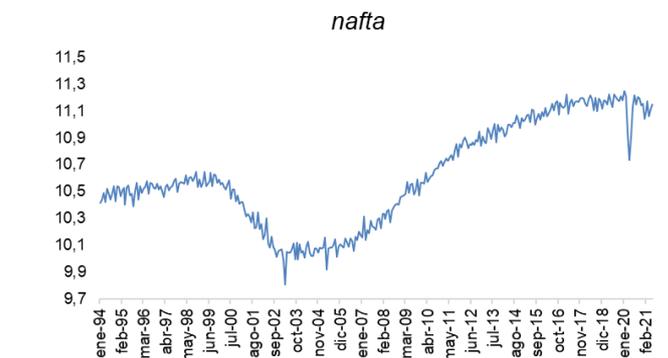
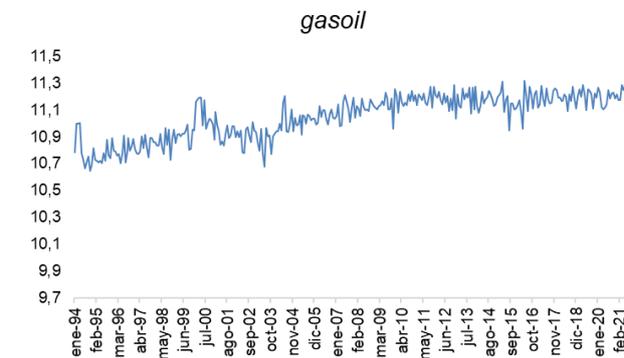
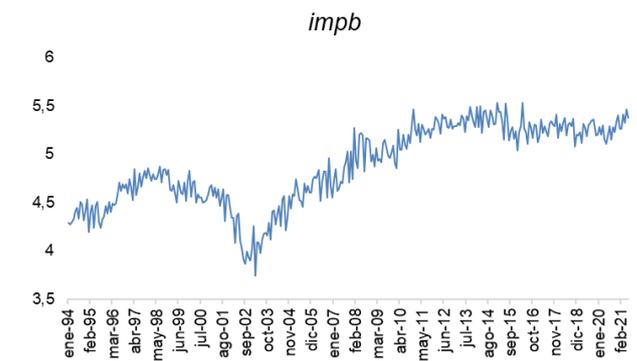
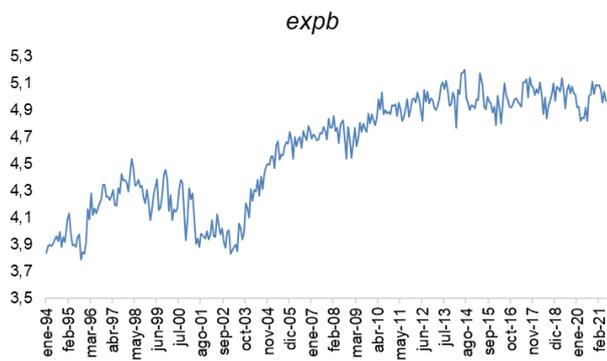
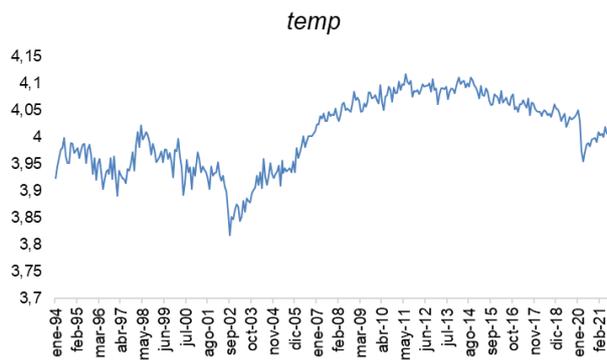
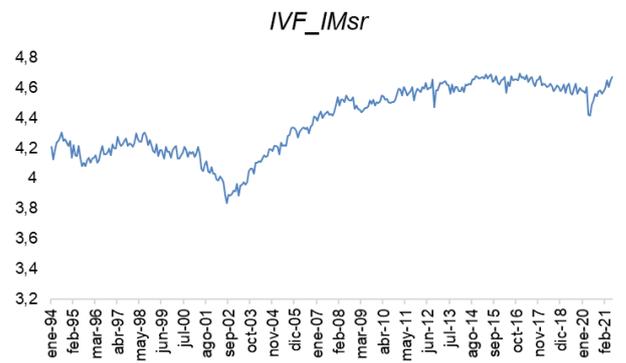
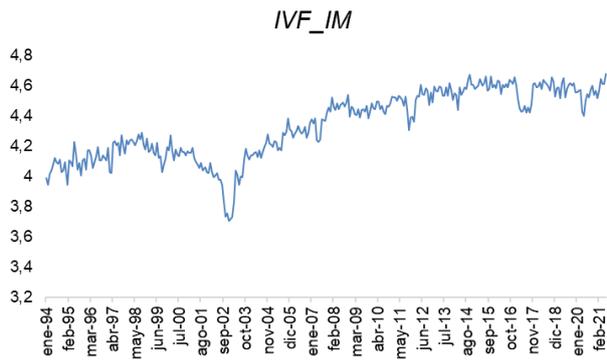
ANEXO 1

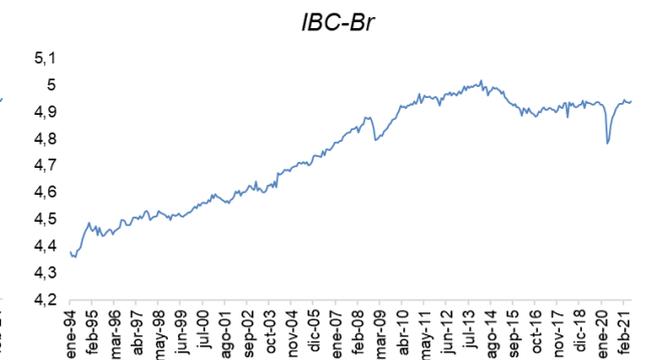
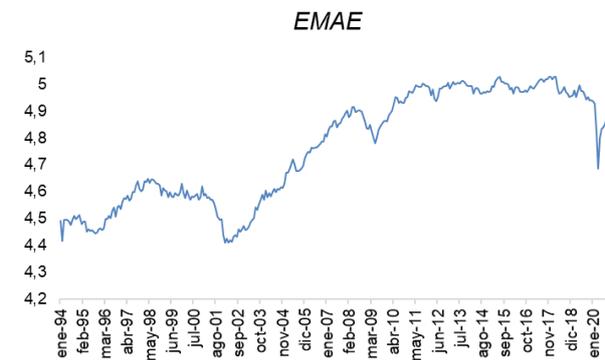
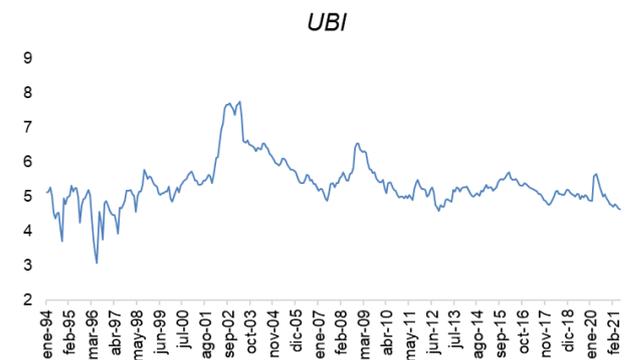
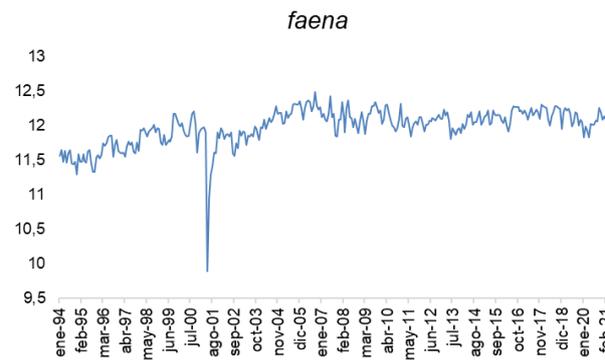
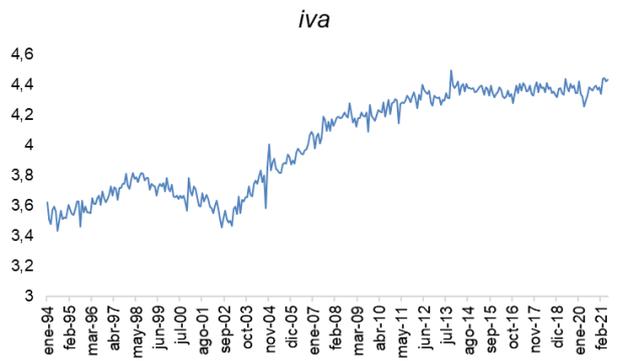
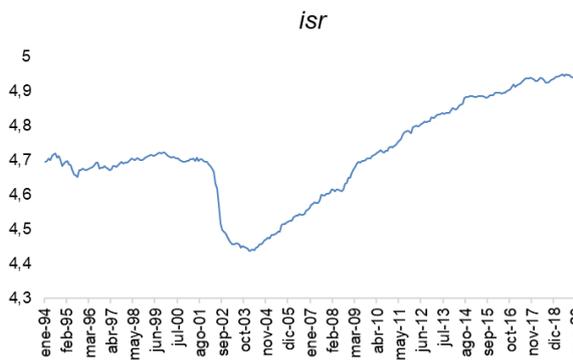
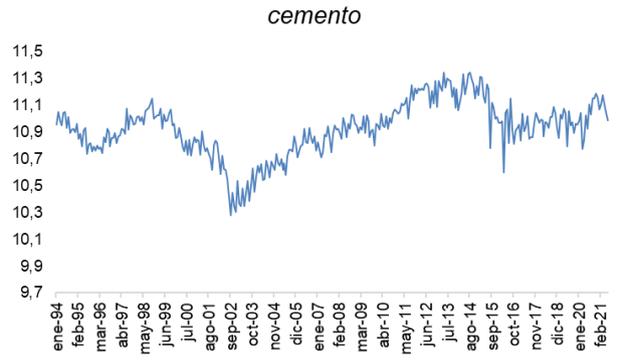
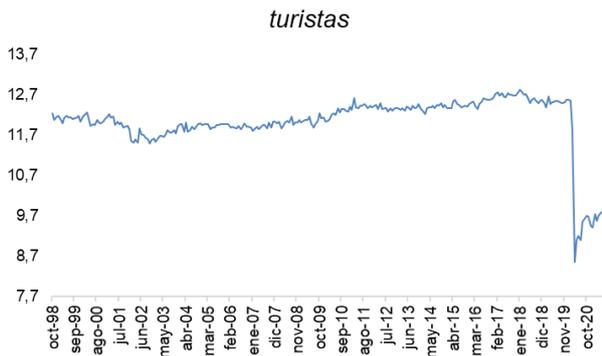
Descripción de las variables

Variable	Forma de llamarla	Fuente	Frecuencia	Rezagó	Gets_E	Gets_D	Lasso_E	Lasso_D	DF_A	DF_R	EP_BR
Índice de volumen Físico del Producto Interno Bruto a precios de 2016	<i>PIB trimestral</i>	BCU	Trimestral	75	X	X	X	X	X	X	X
Índice de volumen Físico del Indicador Mensual de la Industria Manufacturera	<i>IVF_IM</i>	INE	Mensual	45						X	X
Índice de volumen Físico del Indicador Mensual de la Industria Manufacturera sin refinería	<i>IVF_IMsr</i>	INE	Mensual	45	X	X	X	X	X		
Tasa de Empleo	<i>temp</i>	econ.uy en base a INE	Mensual	30	X	X	X	X	X	X	X
Venta de automóviles 0 kilómetro	<i>vautos</i>	ASCOMA	Mensual	10	X	X	X	X	X	X	
Exportaciones de bienes a precios constantes de 2005	<i>expb</i>	BCU	Mensual	45	X	X	X	X	X	X	X
Importaciones de bienes a precios constantes de 2005	<i>impb</i>	BCU	Mensual	45	X	X	X	X	X	X	
Volumen de ventas de gasoil	<i>gasoil</i>	ANCAP/BCU	Mensual	15	X	X	X	X	X	X	
Volumen de ventas de nafta	<i>nafta</i>	ANCAP/BCU	Mensual	15	X	X	X	X	X	X	
Número de turistas que ingresan al país	<i>turistas</i>	Ministerio de Turismo	Mensual	45		X		X	X	X	
Venta de cemento al mercado interno	<i>cemento</i>	CIU	Mensual	30	X	X	X	X	X	X	
Índice de salario real	<i>isr</i>	econuy en base a INE	Mensual	30	X	X	X	X	X		
Recaudación real del Impuesto al Valor Agregado (IVA)	<i>iva</i>	econ.uy en base a DGI	Mensual	30	X	X	X	X	X		
Cabezas faenadas de bovinos y ovino	<i>faena</i>	INAC	Mensual	10	X	X	X	X	X		
Índice de Actividad Económica del Banco Central de Brasil	<i>IBC-Br</i>	BCR	Mensual	50	X	X	X	X	X		
Indicador Mensual de Actividad Económica para Argentina	<i>EMAE</i>	INDEC	Mensual	50	X	X	X	X	X		
Uruguay Bond Index	<i>UBI</i>	República-AFAP	Mensual	1	X	X	X	X	X		

Nota: El rezago es aproximado, y se presenta como el número de días en la que la información está disponible, una vez terminado el periodo por el que informan. Las últimas cinco columnas señalan qué variables son incluidas en cada uno de los modelos para los que se presenta resultados.

Gráficos de las series mensuales desestacionalizadas (en logaritmos)





Modelos para desestacionalización y correlación respecto al PIB

Serie	Modelo	Coeficiente de Correlación respecto al PIB	
		Niveles	Primera diferencia
IVF_IM	(2 1 2)(0 1 1)	0,95	0,52
IVF_IMsr	(0 1 1)(0 1 1)	0,95	0,71
temp	(0 1 1)(0 1 1)	0,80	0,41
vautos	(2 1 0)(0 1 1)	0,78	0,55
expb	(1 1 1)(0 1 1)	0,94	0,19
impb	(2 1 1)(0 1 1)	0,95	0,38
gasoil	(0 1 1)(0 1 1)	0,83	0,03
nafta	(2 1 1)(0 1 1)	0,92	0,60
turistas	(0 1 1)(0 1 1)	0,58	0,64
cemento	(0 1 1)(1 1 1)	0,68	0,37
isr	(0 1 0)(0 1 1)	0,88	0,41
iva	(0 1 1)(0 1 1)	0,96	0,29
faena	(0 1 2)(0 1 1)	0,55	0,18
IBC-Br	(1 1 0)(0 1 1)	0,89	0,54
EMAE	(0 1 0)(0 1 1)	0,93	0,59
UBI	(0 1 1)	-0,34	-0,22

Nota: En la columna Modelos se presenta la representación SARIMA (p d q) (sp sd sq) utilizada para la desestacionalización de las series, siendo p y q los componentes autorregresivos y de medias móviles del componente regular; d las diferencias regulares; sp y sq los componentes autorregresivos y de medias móviles de la parte estacional; mientras que sd se corresponde con las diferencias estacionales. Los coeficientes de correlación son calculados sobre las series desestacionalizadas en frecuencia trimestral. Se presenta el coeficiente Pearson

ANEXO 2

Se detalla de forma matricial la representación Estado Espacio utilizada para estimar los modelos de Factores Dinámicos en frecuencias mixtas para el caso en el que el número de factores es igual a r y estos factores al igual que los componentes idiosincráticos siguen una representación autorregresiva de orden 1.

$$x_t = \tilde{\Lambda}s_t + \eta_t, \text{ Ecuación de Medida (1)}$$

$$s_t = \tilde{A}s_{t-1} + v_t, \text{ Ecuación de estado (2)}$$

$$\begin{bmatrix} y_t^M \\ y_t^Q \end{bmatrix} = \begin{bmatrix} \Lambda^M & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3}\Lambda^Q & \frac{2}{3}\Lambda^Q & \Lambda^Q & \frac{2}{3}\Lambda^Q & \frac{1}{3}\Lambda^Q & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \epsilon_t^M \\ e_t^Q \\ e_{t-1}^Q \\ e_{t-2}^Q \\ e_{t-3}^Q \\ e_{t-4}^Q \end{bmatrix} + \begin{bmatrix} \epsilon_t^M \\ \epsilon_t^Q \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ \epsilon_t^M \\ e_t^Q \\ e_{t-1}^Q \\ e_{t-2}^Q \\ e_{t-3}^Q \\ e_{t-4}^Q \end{bmatrix} = \begin{bmatrix} A_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_r & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_M & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \alpha_Q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \\ f_{t-5} \\ \epsilon_{t-1}^M \\ e_{t-1}^Q \\ e_{t-2}^Q \\ e_{t-3}^Q \\ e_{t-4}^Q \\ e_{t-5}^Q \end{bmatrix} + \begin{bmatrix} u_t \\ 0 \\ 0 \\ 0 \\ 0 \\ \epsilon_t^M \\ \epsilon_t^Q \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (2)$$

Donde y_t^M es un vector de dimensión $N_m * 1$ con las diferencias logarítmicas variables mensuales, y_t^Q un escalar que contiene la diferencia del logaritmo del PIB trimestral; f_t un vector de dimensión $r * 1$ que contiene los r factores inobservables. ϵ_t^M es un

vector de $N_m * 1$ que contiene los componentes idiosincráticos propios de cada una de las variables mensuales. ϵ_t^Q es el componente idiosincrático del PIB trimestral, al tiempo que e_t^Q es el componente idiosincrático de la contraparte mensual - inobservable- del PIB trimestral.

I_r son matrices identidad de dimensión rxr , Λ^M es un vector de N_m coeficientes que relaciona las variables mensuales con los factores en frecuencia mensual.

Λ^Q es una matriz que contiene los coeficientes que relacionan el PIB trimestral con los factores en frecuencia mensual de forma contemporánea y respecto sus cuatro primeros cuatro rezagos.

α_M es una matriz diagonal con N_m elementos en la misma, que se corresponden con los coeficientes autorregresivos que modelan el componente idiosincrático de las variables mensuales.

α_Q es el coeficiente autorregresivo que modela el componente mensual idiosincrático del PIB trimestral

A_1 es una matriz de rxr con los coeficientes autorregresivos de orden 1 para los factores y u_t el componente idiosincrático de la modelización de los factores.