

2024

MODELO PARA LA PREVISIÓN DEL PIB
DE LA ECONOMÍA ESPAÑOLA
A CORTO PLAZO EN TIEMPO REAL
(SPAIN-STING): NUEVA ESPECIFICACIÓN
Y REEVALUACIÓN DE SU CAPACIDAD
PREDICTIVA

BANCO DE **ESPAÑA**
Eurosistema

Documentos Ocasionales
N.º 2406

Ana Gómez Loscos, Miguel Ángel González Simón
y Matías José Pacce

**MODELO PARA LA PREVISIÓN DEL PIB DE LA ECONOMÍA ESPAÑOLA A CORTO PLAZO
EN TIEMPO REAL (SPAIN-STING): NUEVA ESPECIFICACIÓN Y REEVALUACIÓN DE SU
CAPACIDAD PREDICTIVA**

**MODELO PARA LA PREVISIÓN DEL PIB DE LA ECONOMÍA
ESPAÑOLA A CORTO PLAZO EN TIEMPO REAL
(SPAIN-STING): NUEVA ESPECIFICACIÓN Y REEVALUACIÓN
DE SU CAPACIDAD PREDICTIVA (*)**

Ana Gómez Loscos

BANCO DE ESPAÑA

Miguel Ángel González Simón

BANCO DE ESPAÑA

Matías José Pacce

BANCO DE ESPAÑA

(*) Este documento ha sido elaborado por Ana Gómez Loscos, Miguel Ángel González Simón y Matías José Pacce, de la Dirección General de Economía y Estadística.

Documentos Ocasionales. N.º 2406

Marzo 2024

<https://doi.org/10.53479/36136>

La serie de Documentos Ocasionales tiene como objetivo la difusión de trabajos realizados en el Banco de España, en el ámbito de sus competencias, que se consideran de interés general.

Las opiniones y análisis que aparecen en la serie de Documentos Ocasionales son responsabilidad de los autores y, por tanto, no necesariamente coinciden con los del Banco de España o los del Eurosistema.

El Banco de España difunde sus informes más importantes y la mayoría de sus publicaciones a través de la red Internet en la dirección <http://www.bde.es>.

Se permite la reproducción para fines docentes o sin ánimo de lucro, siempre que se cite la fuente.

© BANCO DE ESPAÑA, Madrid, 2024

ISSN: 1696-2230 (edición electrónica)

Resumen

El incremento de la volatilidad observada en la mayoría de los indicadores económicos tras la irrupción del COVID-19 redujo la capacidad predictiva de los modelos de previsión a corto plazo. En este documento se presenta una revisión del modelo Spain-STING —una de las herramientas que utiliza el Banco de España para la predicción a corto plazo del crecimiento intertrimestral del PIB— al objeto de mejorar su capacidad predictiva tras la pandemia. En particular, la revisión comporta tres cambios principales: i) la relación entre los indicadores incluidos en el modelo y el componente común estimado pasa a ser contemporánea para todos los indicadores, en lugar de adelantada en el tiempo para alguno de ellos; ii) se permite que la varianza del componente común estimado pueda sufrir cambios en el tiempo, al modelarse a través de un proceso estocástico; iii) se revisa el conjunto de indicadores con el fin de incluir solo aquellos que aportan la información más relevante a la hora de predecir el crecimiento del PIB tras la pandemia. Estas modificaciones redundan en una mejora sustancial de la capacidad predictiva de Spain-STING en el período posterior a la pandemia y mantienen la correspondiente al período anterior a ella.

Palabras clave: ciclos económicos, economía española, modelos de factores dinámicos, COVID-19.

Códigos JEL: C22, E27, E32.

Abstract

After the outbreak of the COVID-19 pandemic, most economic indicators experienced an increase in the observed volatility, which reduced the accuracy of nowcasting models. In this paper, we present a revision of the Spain-STING model – one of the tools used by the Banco de España to nowcast the quarterly GDP growth rate – that focuses on improving the model's predictive ability after the pandemic. Specifically, three main changes are made: (i) the relationship between the indicators and the estimated common factor is now contemporaneous, and not leading for some of the indicators; (ii) the variance of the common component is estimated by a stochastic process to allow it to vary over time; (iii) the set of variables is revised with the aim of including only those that add the most relevant information to the nowcast of the quarterly GDP growth rate after the pandemic. All three modifications imply a notable improvement in the Spain-STING nowcasting performance during the post-pandemic period, while maintaining its pre-pandemic accuracy.

Keywords: business cycles, Spanish economy, dynamic factor models, COVID-19.

JEL classification: C22, E27, E32.

Índice

Resumen 5

Abstract 6

1 Introducción 8

2 El modelo Spain-STING anterior a la pandemia 10

2.1 Descripción del modelo 10

2.2 Capacidad predictiva del modelo hasta 2019 y deterioro de esta a raíz de la inclusión del período de la pandemia 12

3 El nuevo modelo Spain-STING 16

3.1 Modificación de la relación temporal entre las variables y el componente común 16

3.2 Incorporación de volatilidad estocástica asociada al factor común 19

3.3 Modificación del conjunto de variables del modelo 21

3.4 Evaluación de la capacidad predictiva del modelo *Revisado* frente al *Anterior* 24

4 Consideraciones finales 28

Bibliografía 29

Anejo 1. Aproximación de vectores apilados para la representación del modelo 31

Anejo 2. Estimación bayesiana 32

Anejo 3. Período COVID como no observado 33

1 Introducción

El modelo Spain-STING¹ es una de las herramientas que utiliza el Banco de España para la predicción del crecimiento intertrimestral del PIB a corto plazo. Spain-STING utiliza un conjunto de indicadores de actividad económica —mensuales y trimestrales—, y descompone la evolución de sus series temporales en un factor común y un componente idiosincrásico. El factor común captura las dinámicas comunes de los distintos indicadores, mientras que el componente idiosincrásico refleja la parte de la evolución de cada indicador que no puede atribuirse a dicho componente común.

Hasta diciembre de 2019, Spain-STING mostró una capacidad predictiva notable del crecimiento del PIB de la economía española. Sin embargo, la inclusión del período de la pandemia del COVID-19 ha incrementado los errores de previsión de las herramientas de modelización de componentes inobservables y, en particular, de este modelo. Esto se debe, fundamentalmente, a los intensos cambios observados en las dinámicas de las variables y al aumento de su volatilidad, lo cual se habría traducido en una alteración de la relación a largo plazo entre los indicadores y el factor común estimado por el modelo.

En este documento, se plantea una revisión del modelo en tres aspectos fundamentales que permiten acomodar los cambios observados en las dinámicas de las variables como consecuencia de la pandemia. El primero es la reevaluación de la relación temporal que se asume entre las variables incluidas en el modelo y el factor común estimado, que potencialmente puede ser contemporánea, retardada o adelantada. El segundo es la incorporación de volatilidad estocástica para tener en cuenta la mayor variabilidad de las variables durante períodos como el de la pandemia. El tercero es la revisión y modificación del conjunto de indicadores, tanto cuantitativos como cualitativos, que se incluye en el modelo.

Cada una de las modificaciones realizadas se evalúa con el objetivo de mejorar la capacidad predictiva, en el trimestre corriente, del modelo Spain-STING durante el período posterior a la fase más aguda de la pandemia y, a la vez, sin deteriorarla en el período anterior a ella —no se evalúa la capacidad predictiva en situaciones de elevada incertidumbre asociada a escenarios de alta volatilidad, como el observado durante la fase más aguda de la pandemia en 2020—. Es decir, la revisión del modelo se centra en la reducción de los errores de las previsiones durante el período que abarca desde el primer trimestre de 2021 hasta el segundo trimestre de 2023. En este sentido, la monitorización y la revisión de los modelos de predicción resultan necesarias para garantizar la fiabilidad de las previsiones en la cuales basan sus decisiones los agentes económicos, especialmente, tras la irrupción de crisis o de eventos extremos.

¹ La primera versión del modelo Spain-STING se desarrolla en Camacho y Pérez-Quirós (2011). Posteriormente, Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020) extendieron el modelo para predecir conjuntamente tanto el PIB como los componentes de demanda.

Tras esta introducción, el documento se estructura como sigue. En el segundo epígrafe, se describe el marco teórico y metodológico del modelo empleado hasta diciembre de 2019, así como la evolución de su capacidad predictiva antes y después de la pandemia. En el tercer epígrafe, se presentan los tres cambios, mencionados anteriormente, introducidos en el modelo a fin de mejorar su capacidad predictiva. Finalmente, se analiza la efectividad de estos cambios de forma conjunta en términos de la capacidad predictiva del modelo revisado en comparación con el modelo antiguo. En el epígrafe último se exponen las consideraciones finales.

2 El modelo Spain-STING anterior a la pandemia

2.1 Descripción del modelo

El modelo Spain-STING, propuesto inicialmente por Camacho y Pérez-Quirós (2011) y actualizado por Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020), es un modelo de factores dinámicos de pequeña escala que permite el uso de frecuencias mixtas (en particular, la combinación de indicadores mensuales y trimestrales de actividad) y que se utiliza fundamentalmente para la predicción, en tiempo real, del crecimiento trimestral del PIB de la economía española.

Así, la evolución de las series temporales se caracteriza como la suma de dos componentes ortogonales. Concretamente, la tasa de crecimiento de una variable de frecuencia mensual (z_t^j) —o de una variable de frecuencia trimestral (x_t^j)— se expresa como la suma de un factor común (f_t), que recoge las dinámicas comunes de los distintos indicadores incorporados al modelo, y de un componente idiosincrásico (ε_t^j), que refleja la parte de la dinámica de cada indicador que no puede atribuirse al componente común, donde $t = 1, \dots, T$ representa el período temporal expresado en meses y $j = 1, \dots, J$ las variables incluidas en el modelo. Dado el uso de frecuencias mixtas, se hace necesario adecuar el modelo para la combinación de variables expresadas en tasas de crecimiento intermensuales e intertrimestrales. Para ello, se sigue la metodología propuesta por Mariano y Murasawa (2003), según la cual es posible aproximar la tasa intertrimestral de crecimiento de una variable como la media ponderada de sus tasas de crecimiento intermensuales². Por tanto, para una especificación que contiene únicamente una variable trimestral y una variable mensual ($j = 1, 2$), el modelo se describe de la siguiente manera:

$$\begin{pmatrix} x_t^1 \\ z_t^2 \end{pmatrix} = \begin{bmatrix} 1/3\beta_1 & 2/3\beta_1 & \beta_1 & 2/3\beta_1 & 1/3\beta_1 \\ \beta_2 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_t \\ f_{t-1} \\ f_{t-2} \\ f_{t-3} \\ f_{t-4} \end{bmatrix} + \begin{bmatrix} 1/3 & 2/3 & 1 & 2/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_t^1 \\ u_{t-1}^1 \\ u_{t-2}^1 \\ u_{t-3}^1 \\ u_{t-4}^1 \\ u_t^2 \end{bmatrix} \quad (1)$$

$$\phi_f(L) f_t = \varepsilon_t \quad (2)$$

$$\phi_j(L) u_t^j = \varepsilon_t^j \quad (3)$$

donde $\phi_j(L)$ y $\phi_f(L)$ son polinomios de retardos de orden p_j y q , respectivamente, y se asume que los errores se distribuyen como $\varepsilon_t^j \sim N(0, \sigma_j^2)$ y $\varepsilon_t \sim N(0, \sigma_f^2)$ y son independientes entre sí. Los parámetros β_j se denominan cargas factoriales (o factores de carga) y capturan la correlación entre el factor común y las variables.

Este modelo puede ser representado en forma de espacio de los estados y, empleando el filtro de Kalman (véase, por ejemplo, Hamilton, 1994), puede estimarse utilizando un estimador

² La tasa de crecimiento intertrimestral de una variable (x_t) puede aproximarse como la suma de las tasas de crecimiento intermensuales (z_t) de la misma variable mediante la siguiente fórmula: $x_t = (1/3)z_t + (2/3)z_{t-1} + z_{t-2} + (2/3)z_{t-3} + (1/3)z_{t-4}$

Cuadro 1

INDICADORES DEL MODELO SPAIN-STING UTILIZADOS EN ARENCIBIA PAREJA, GÓMEZ LOSCOS, DE LUIS LÓPEZ Y PÉREZ-QUIRÓS (2020)

Indicador	Tipo de indicador	Fuente	Periodicidad	Relación	Fecha de inicio	Retraso en su publicación
Tasa de variación del PIB	Actividad	INE	Trimestral	Contemporánea	Marzo 1990	+ 30 días
ISE sin consumidores	Basado en encuesta	Comisión Europea	Mensual	Adelantada (3 meses)	Enero 1990	0 días
PMI compuesto	Basado en encuesta	IHS Markit	Mensual	Adelantada (3 meses)	Agosto 1990	+ 5 días
Consumo eléctrico	Actividad	Red Eléctrica de España	Mensual	Contemporánea	Febrero 1990	+ 1 días
Afiliados a la Seguridad Social	Actividad	Seguridad Social	Mensual	Contemporánea	Enero 1990	+ 3 días
Ventas de grandes empresas	Actividad	Agencia Tributaria	Mensual	Contemporánea	Febrero 1996	+ 10 días
IPI no energético	Actividad	INE	Mensual	Contemporánea	Febrero 1992	+ 36 días
IPI de construcción (a)	Actividad	INE	Mensual	Adelantada (3 meses)	Febrero 1992	+ 36 días
Crédito a empresas no financieras	Actividad	Banco de España	Mensual	Contemporánea	Febrero 1995	+ 30 días
Importaciones reales de bienes	Actividad	Departamento de Aduanas	Mensual	Contemporánea	Febrero 1991	+ 50 días
Exportaciones reales de bienes	Actividad	y MINECO	Mensual	Contemporánea	Febrero 1991	+ 50 días

FUENTE: Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020).

a En el modelo de Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020) se emplea la variable de Consumo aparente de cemento. Sin embargo, a finales de 2017 un problema con la regularidad en la publicación de esta serie llevó a su sustitución por el IPI de construcción, cuyo contenido informativo puede asimilarse al de la serie de Consumo aparente de cemento.

de máxima verosimilitud. Siguiendo la metodología propuesta por Mariano y Murasawa (2003), la estimación puede adaptarse para incluir observaciones faltantes, lo cual es de especial utilidad, ya que, por un lado, implica que no es necesario disponer de una muestra balanceada al final del período y, por otro, permite tratar el hecho de que las variables trimestrales se observan solo una vez en el trimestre. Para una explicación más detallada del modelo y su estimación, véase Camacho y Pérez-Quirós (2011) y Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020).

El conjunto de indicadores incluido, tras la actualización del modelo por Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020), está compuesto por 11 variables: una de frecuencia trimestral (el PIB) y diez de frecuencia mensual (véase cuadro 1). Las variables de frecuencia mensual pueden dividirse, a su vez, en indicadores de actividad (comúnmente denominados *hard*) e indicadores basados en encuestas (habitualmente denominados *soft*). Cabe mencionar que los indicadores *hard* se incluyen en el modelo en tasas de crecimiento intermensuales de la forma descrita en la ecuación (1). En el caso de los indicadores *soft*, son incluidos en niveles de acuerdo con la siguiente especificación:

$$x_t^i = \sum_{i=0}^{11} \beta_i f_{t-i} + u_t^i \quad (4)$$

donde u_t^i sigue la dinámica descrita en la ecuación (3)³.

Adicionalmente, es de especial importancia la relación que modeliza la dinámica de las variables con la del factor común. En este sentido, existe la posibilidad de que las variables actúen como indicadores contemporáneos —y se relacionen de la forma descrita en la ecuación (1)—, adelantados (x_{jt} se relaciona con f_{t+i}) o retardados (x_{jt} se relaciona con f_{t-i}). Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020) encontraron que la especificación que mejor predecía la tasa trimestral de crecimiento del PIB español (con datos hasta el tercer trimestre de 2016) era la que incorporaba tres de los indicadores mensuales de forma adelantada en un trimestre⁴. En concreto, los dos indicadores *soft* y el IPI de construcción.

En el cuadro 1 se especifican, por un lado, los indicadores incluidos en el modelo —PIB, Indicador de Sentimiento Económico (ISE) sin consumidores, Purchasing Managers Index (PMI) compuesto, Consumo eléctrico, Afiliados a la Seguridad Social, Ventas de grandes empresas, Índice de Producción Industrial (IPI) no energético, IPI de construcción, Crédito a empresas no financieras, Importaciones reales de bienes y Exportaciones reales de bienes— y, por otro lado, se detalla la periodicidad de cada indicador, su tipo (*hard* o *soft*), el mes de comienzo de la muestra, el retraso en su publicación y cuál es la relación con el factor común que se asume en el modelo (contemporáneo o adelantado).

2.2 Capacidad predictiva del modelo hasta 2019 y deterioro de esta a raíz de la inclusión del período de la pandemia

El modelo especificado de la forma descrita en las ecuaciones (1) a (4) —en adelante, «modelo *Anterior*»— mostraba, hasta diciembre de 2019, una capacidad predictiva notable para el PIB de la economía española. En el gráfico 1.1 se presenta la evolución de la previsión del PIB —obtenida con la información disponible a mediados del tercer mes de cada uno de los trimestres para los cuales se realiza la previsión— en comparación tanto con la primera (estimación *flash*) como con la segunda estimación del crecimiento intertrimestral del PIB, publicada por el Instituto Nacional de Estadística (INE), para cada uno de los trimestres entre 2015 y 2019. Complementariamente, en el gráfico 1.3, se muestran los errores absolutos de previsión correspondientes al gráfico 1.1. Como puede observarse, la capacidad predictiva de las previsiones se mantiene estable y en niveles reducidos a lo largo de todos los años, situándose el error medio absoluto (MAE, por sus siglas en inglés) del período en torno a 0,1 puntos porcentuales (pp). Estos resultados sugieren que el modelo *Anterior* tenía una capacidad predictiva elevada hasta finales de 2019.

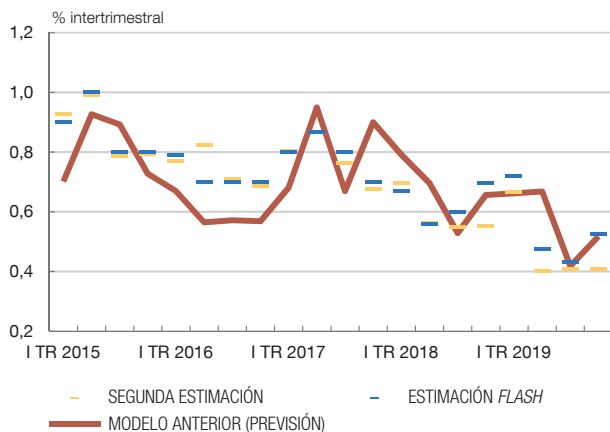
3 Tal y como se expone en Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020), al incorporar los indicadores en niveles puede darse un problema potencial respecto a la modelización por considerarse de forma simultánea variables estacionarias e integradas. Para resolver este problema, se sigue la indicación de la Comisión Europea (2006) según la cual los indicadores *soft* se relacionan con la tasa interanual de crecimiento de la variable de interés. Por lo tanto, se establece que la fuente de su raíz unitaria provenga de que el nivel de estas variables depende de una media móvil de 12 meses del factor común.

4 Esto significa que, si el indicador mensual fuese de tipo *hard*, la relación establecida en la ecuación (1) se describiría como $x_t^i = \beta_i f_{t+3} + u_t^i$, mientras que, si la variable fuese *soft*, la relación se describiría como $x_t^i = \sum_{i=0}^{11} \beta_i f_{t+3-i} + u_t^i$

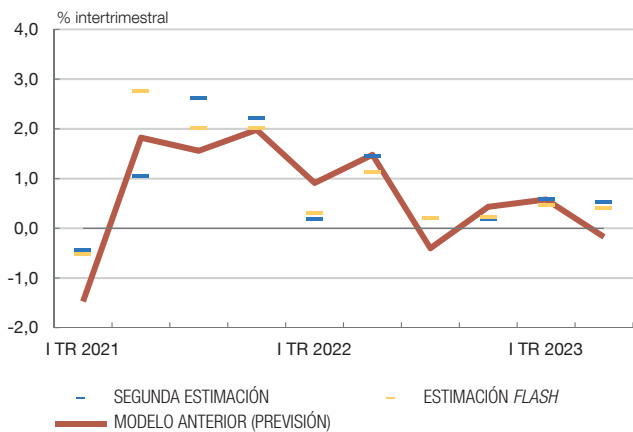
Gráfico 1

ERRORES DE PREVISIÓN DEL MODELO ANTERIOR ANTES Y DESPUÉS DE LA INCLUSIÓN DEL PERÍODO POS-COVID

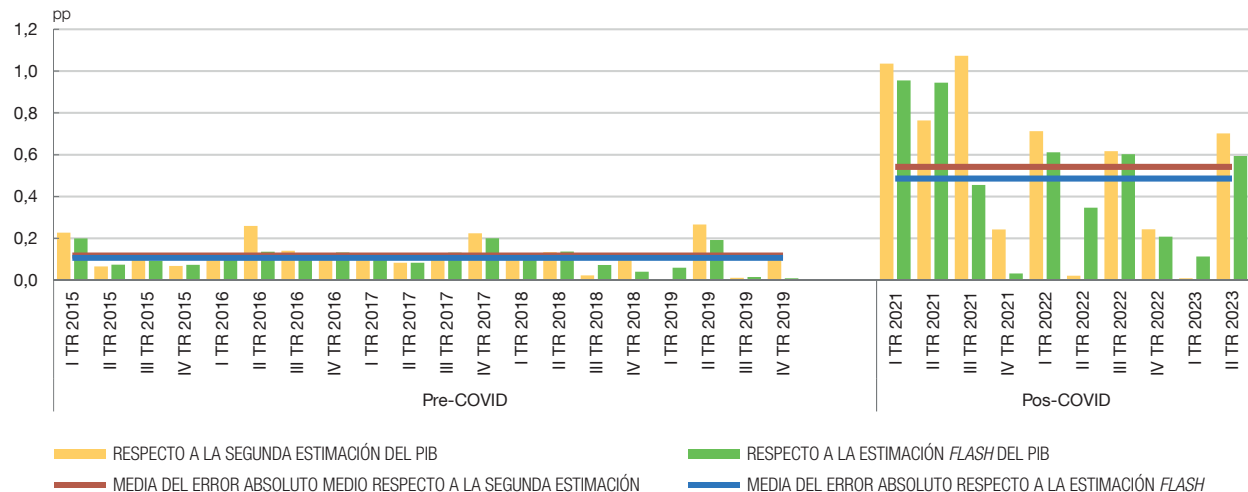
1 PIB: OBSERVADO Y PREVISIONES. PERÍODO 2015-2019



2 PIB: OBSERVADO Y PREVISIONES. PERÍODO 2021-2023



3 ERRORES ABSOLUTOS EN CADA TRIMESTRE. MODELO ANTERIOR. PERÍODO PRE-COVID (I TR 2015 - IV TR 2019) Y PERÍODO POS-COVID (I TR 2021 - II TR 2023)



FUENTE: Elaboración propia.

NOTA: Las líneas horizontales representan la media de los errores absolutos en el período pre-COVID y pos-COVID, de forma separada, frente a las del PIB y la estimación *flash*.

Sin embargo, la inclusión del período posterior a la pandemia del COVID-19⁵ incrementó de forma notable los errores de las previsiones del modelo Spain-STING, al igual que en el caso de la mayoría de los modelos de previsión basados en factores inobservables, debido a la complejidad en el tratamiento de las variaciones experimentadas por las variables económicas, que aumentó de forma significativa durante la pandemia. De este modo, los gráficos 1.2 y 1.3 muestran el fuerte deterioro de las previsiones del modelo, reflejado en un incremento notable de los errores, hasta el entorno de los 0,5 pp, durante el período pos-COVID (I TR 2021 - II TR 2023).

5 En el presente análisis no se toma en cuenta el período de la fase más aguda del COVID-19, ya que, para su tratamiento, se necesita un enfoque metodológico diferente, que se aleja del objetivo planteado en este documento.

Cuadro 2

VARIANZA OBSERVADA POR PERÍODO

	Total (2015-2023)	Pre-COVID (2015-2019)	Pos-COVID (2021-2023)
Tasa de variación del PIB	20,93	0,07	1,63
ISE sin consumidores	48,45	5,12	19,96
PMI compuesto	47,12	3,53	24,36
Consumo eléctrico en industria	4,86	3,18	1,81
Afiliados a la Seguridad Social	4,64	0,02	0,18
Ventas de grandes empresas	12,83	0,52	4,78
IPI no energético	18,45	2,45	1,27
IPI de construcción	82,47	2,11	11,42
Crédito a empresas no financieras	1,36	0,65	1,73
Importaciones reales de bienes	18,96	7,15	7,29
Exportaciones reales de bienes	26,24	6,48	11,68

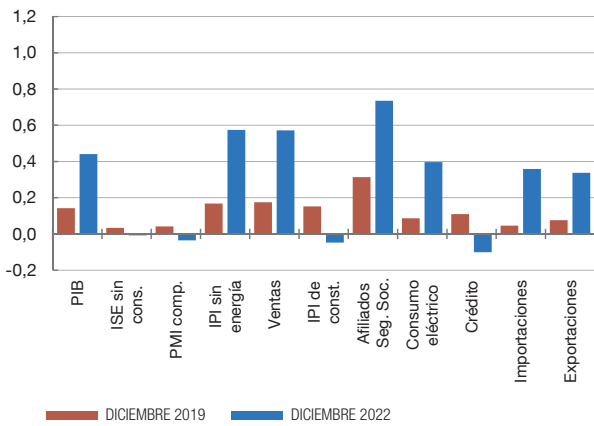
FUENTES: Elaboración propia.

Una posible causa que explica el deterioro de la capacidad predictiva del modelo puede encontrarse en la elevada variabilidad que mostraron las variables a partir del año 2020. El cuadro 2 presenta la varianza de cada una de las variables incluidas en Spain-STING durante el período previo y posterior al COVID-19, así como para el conjunto de la muestra. Como puede apreciarse, la volatilidad de los indicadores ha experimentado, en general, un notable aumento en el período posterior a la pandemia.

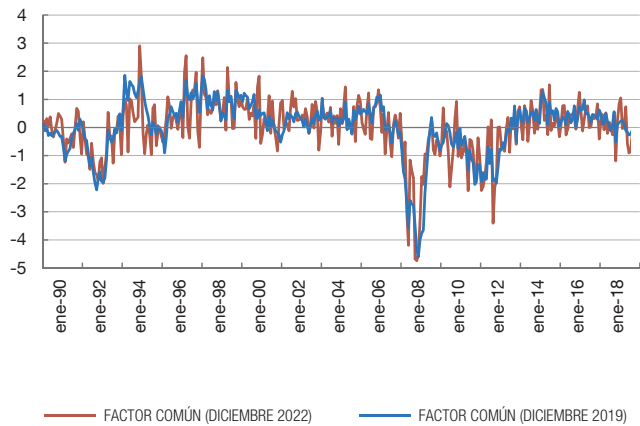
Además, el efecto de la pandemia sobre la dinámica de las variables y, especialmente, el incremento de su volatilidad, habría afectado de forma directa a la relación de largo plazo existente entre los distintos indicadores y, por tanto, a la dinámica del componente común que se extrae mediante el modelo Spain-STING —véase ecuación (1)—. A modo de ejemplo, se muestran dos métricas que permiten visualizar cómo el período del COVID-19 afectó a las estimaciones del modelo. Por un lado, en el gráfico 2.1, se observa, para distintos indicadores, un cambio sustancial en los factores de carga (β_j) estimados por el modelo *Anterior* con información disponible hasta diciembre de 2022 (incluyendo el período COVID-19) frente a los obtenidos con información hasta diciembre de 2019 (período pre-COVID-19). En algunos casos particulares, como el del ISE sin consumidores, el PMI compuesto, el IPI de construcción y el crédito a sociedades no financieras, se produce un cambio de signo de los factores de carga, lo que indica un cambio en el sentido de la relación de estas variables con el componente común. Por otro lado, en el gráfico 2.2, se muestra el componente común estimado por el modelo *Anterior* con información disponible, nuevamente, hasta diciembre de 2019 o hasta diciembre de 2022, y ponderado por la respectiva desviación estándar de cada componente en el período comprendido entre 1990 y 2019. Como puede apreciarse, la tasa de variación intermensual del factor común se incrementa significativamente con la inclusión del período del COVID-19 (línea roja). En términos de variación intermensual absoluta, la tasa del factor común estimado incluyendo

EFECTO DE LA INCLUSIÓN DEL PERÍODO POS-COVID EN LAS ESTIMACIONES DEL MODELO ANTERIOR

1 FACTORES DE CARGA ESTIMADOS CON INFORMACIÓN HASTA DICIEMBRE



2 FACTOR COMÚN ESTIMADO CON INFORMACIÓN HASTA DICIEMBRE DE 2019 O DICIEMBRE DE 2022



FUENTE: Elaboración propia.

NOTA: En ambos gráficos se muestran los resultados de las estimaciones que resultan del modelo *Anterior*, incluyendo la información disponible hasta diciembre de 2019 o hasta diciembre de 2022. Los factores de carga se refieren a los parámetros β de la ecuación (1). Cada factor común en el gráfico 2.2 se encuentra dividido por la desviación estándar del propio factor durante el período 1990-2019.

el período de la pandemia triplica, en promedio, a la tasa estimada sin incluir dicho período. El incremento de la variabilidad intermensual del factor estimado refleja que las relaciones de largo plazo de las variables no son capturadas de igual forma una vez que se incluye el período del COVID-19 en la estimación.

3 El nuevo modelo Spain-STING

3.1 Modificación de la relación temporal entre las variables y el componente común

Como se ha indicado en el epígrafe 2, en el modelo propuesto por Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020), el PMI compuesto, el ISE sin consumidores y el IPI de construcción se introducen en el modelo con un adelanto de tres meses respecto al resto de las variables. La justificación para considerarlas indicadores adelantados provenía del hecho de que las dos primeras se refieren a encuestas que recogen expectativas (los cuestionarios se refieren a la evolución que los agentes esperan tener en un horizonte temporal de tres meses), mientras que el IPI de construcción, por su propia idiosincrasia, podría adelantar el comportamiento de la construcción y, por tanto, el del PIB. No obstante, en línea con la evidencia del epígrafe 2, es posible que la relación entre las mencionadas variables y el PIB se haya visto alterada tras la pandemia.

Al objeto de analizar esta hipótesis, el gráfico 3 muestra, para el período anterior a la pandemia (1990-2019) y para el período que la incluye (1990-2023), la correlación entre la tasa de crecimiento interanual del PIB y el nivel⁶ del PMI compuesto y del ISE sin consumidores con diferentes desfases temporales. Es decir, se estiman diferentes correlaciones entre los indicadores y el PIB, ya sean contemporáneas, con meses de retardo o con meses de adelanto. A modo de ejemplo, en el eje de abscisas, el valor t+1 (correlación adelantada en un mes) presenta la correlación del promedio trimestral del indicador en los meses de febrero, mayo, agosto y noviembre respecto a la tasa interanual del PIB de marzo, junio, septiembre y diciembre, respectivamente.

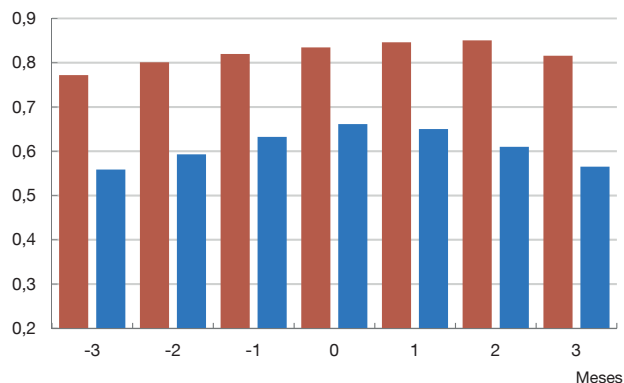
De acuerdo con los resultados obtenidos, para el período 1990-2019, tanto en el caso del ISE sin consumidores como en el del PMI compuesto, la mayor correlación con la tasa interanual del PIB se alcanzaba cuando cada uno de esos indicadores se adelantaba en al menos dos meses (véase gráfico 3). Sin embargo, para el período 1990-2023, una vez incluido el período de la pandemia, estas relaciones cambian: en el caso del ISE sin consumidores la mayor correlación se observa en la relación contemporánea y en el del PMI compuesto, cuando el indicador se adelanta en un mes, si bien este valor solo es ligeramente superior al observado en la correlación contemporánea. Una posible explicación de esta modificación en la relación temporal de estas variables respecto al PIB es la intensa caída de sus valores en los meses de mayor restricción a la movilidad y, por tanto, de mayor caída de la actividad. En el caso del IPI de construcción, el análisis de correlación se realiza entre la tasa de crecimiento intertrimestral de esta variable y la del PIB⁷. Como puede apreciarse en el gráfico 3, la correlación adelantada en tres meses disminuye con intensidad una vez incluido el período posterior a la irrupción de la pandemia en la muestra, siendo la correlación que adelanta al IPI de construcción en un

⁶ Se emplea una media móvil de tres meses.

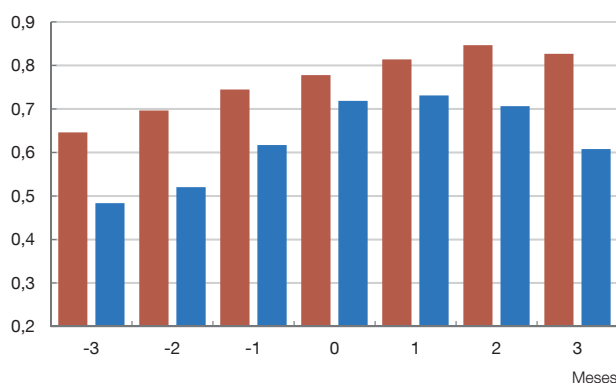
⁷ Para computar la tasa intertrimestral de crecimiento del IPI de construcción se estima una media móvil de tres meses de la variable y se computa la tasa intertrimestral en cada uno de los meses.

CORRELACIONES CRUZADAS ENTRE EL PIB Y DETERMINADAS VARIABLES SELECCIONADAS

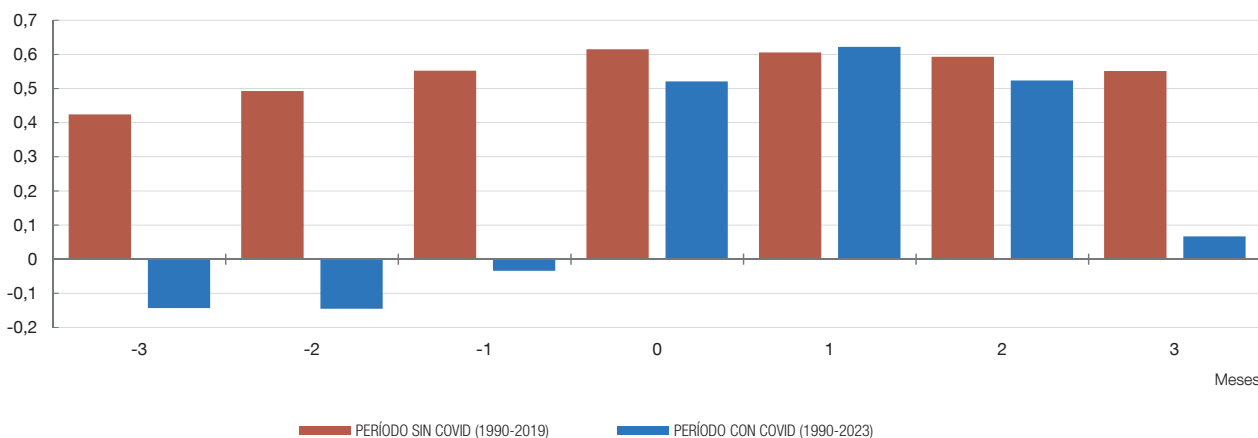
1 CORRELACIÓN CRUZADA DEL ISE SIN CONSUMIDORES FRENTE AL PIB



2 CORRELACIÓN CRUZADA DEL PMI COMPUESTO FRENTE AL PIB



3 CORRELACIÓN CRUZADA DEL IPI DE CONSTRUCCIÓN FRENTE AL PIB



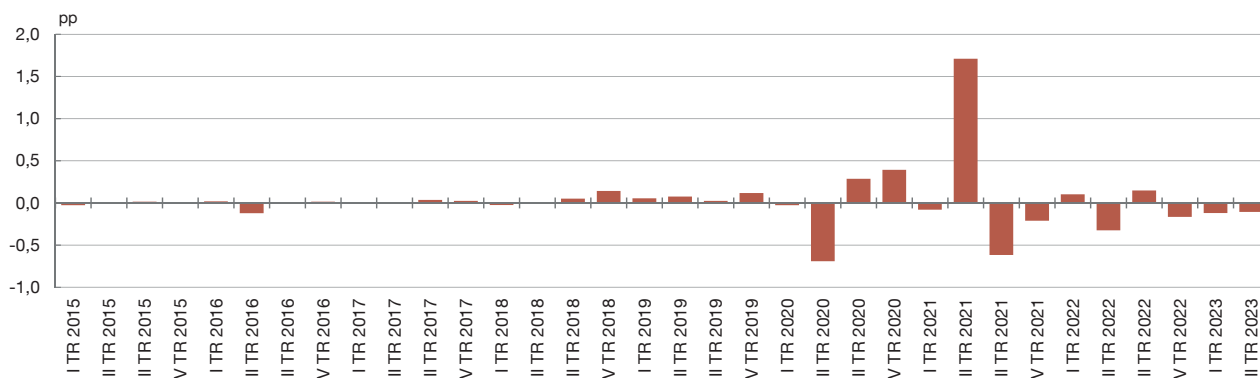
FUENTE: Elaboración propia.

NOTA: En el caso del ISE sin consumidores y del PMI compuesto, se presenta la correlación cruzada respecto a la tasa interanual del PIB. Para el IPI de construcción, se calcula frente a la tasa intertrimestral del PIB. El eje de abscisas representa el número de meses de diferencia en términos de correlación.

mes respecto del PIB la que mayor valor muestra, si bien la correlación contemporánea también es elevada. Los cambios observados en las correlaciones temporales analizadas llevan a plantear la alternativa de que las tres variables puedan haber dejado de ser indicadores adelantados de actividad, por lo que se analiza la alternativa de que se relacionen con el factor común de forma contemporánea⁸. En el caso del resto de las variables, que en el modelo de Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020) entraban con una relación contemporánea, no se observa un cambio, tras la inclusión del período COVID en las correlaciones temporales analizadas.

⁸ La elección de la relación contemporánea y no adelantada en un mes para los casos del PMI compuesto y del IPI de construcción, como sugeriría el análisis de correlación, se basa en la búsqueda de la parsimonia del modelo, ya que el resto de las variables son incorporadas de forma contemporánea y el adelantamiento de estas variables en solo un mes no genera cambios significativos respecto de las predicciones obtenidas.

Gráfico 4

DIFERENCIA ENTRE LA ESTIMACIÓN FLASH Y LA SEGUNDA ESTIMACIÓN DEL PIB

FUENTE: Elaboración propia.

Un aspecto fundamental para valorar las potenciales ganancias de revisar la relación temporal de las variables es estimar el error de previsión obtenido incluyendo o no el cambio en la especificación del modelo. Para ello, en el tercer mes de cada trimestre, desde enero de 2015 hasta junio de 2023, se realiza una previsión del crecimiento intertrimestral del PIB bajo las dos especificaciones descritas. Así, se considera la información que se encontraba disponible el día 23 de cada mes y se realiza la predicción a uno y a dos trimestres. En otras palabras, las estimaciones se realizan en tiempo real⁹. Posteriormente, se computa el error de cada una de las mencionadas estimaciones con respecto a la estimación *flash* del PIB y respecto a la segunda estimación de este. Es importante señalar que, entre ambas estimaciones, en ocasiones, se producen revisiones importantes, las cuales han sido especialmente notables desde el primer trimestre de 2020 hasta el segundo trimestre de 2022 (véase gráfico 4).

Los errores de previsión cometidos en el período posterior a la pandemia no se reducen a partir del cambio propuesto en la especificación de la relación temporal de las tres variables mencionadas y el factor común, si bien tampoco se incrementan. No obstante, el cambio introducido logra una mejora sustancial en términos de la interpretación económica del modelo. Como se ha mencionado en el epígrafe 2.2, la relación histórica entre el componente común, que puede interpretarse como una medida de actividad, y el ISE sin consumidores, el PMI compuesto y el IPI de construcción, que antes del COVID-19 era positiva, pasó a ser negativa (véase gráfico 2.1). Es decir, variaciones positivas en estas variables pasaron a corresponderse con una menor actividad económica. Esta circunstancia, que, desde una perspectiva económica, resulta poco intuitiva, se revierte a partir del cambio asumido en la relación temporal de estas variables con el componente común (es decir, pasa a ser positiva). Por este motivo, se considera que la inclusión de estas tres variables

⁹ La previsión en tiempo real hace referencia a las estimaciones que se realizan con la información de la que disponía el analista en cada uno de los momentos del tiempo.

con una relación contemporánea con el factor común es la adecuada, y en el resto del documento se incluirán de esta manera en el modelo.

3.2 Incorporación de volatilidad estocástica asociada al factor común

Con el objeto de lidiar con el fuerte incremento de la volatilidad de los indicadores que se utilizan en el modelo (véase cuadro 2), se considera la posibilidad de que la volatilidad del factor común estimada por el modelo varíe en el tiempo. En la versión previa del modelo, descrita en el epígrafe anterior, una característica fundamental del comportamiento del factor común era que la varianza asociada a él se asumía invariante en el tiempo. Esta especificación puede resultar excesivamente rígida en escenarios de crisis o eventos extremos, cuando la volatilidad de los indicadores suele incrementarse con intensidad y, por tanto, afectar a la dinámica del componente común estimado. Al permitir que la varianza asociada al factor común varíe en el tiempo, se logra que los períodos en los que la volatilidad de los indicadores es más elevada tengan una menor relevancia para la estimación del componente común, ya que es el incremento de la volatilidad del factor, y no el factor en sí mismo, lo que explicaría una mayor parte de la dinámica común de las variables.

Para introducir volatilidad estocástica en la dinámica del factor común, se reemplaza la ecuación (2) de la versión anterior del modelo por la siguiente especificación:

$$\phi^f(L) f_t = \sigma_{f_t} \epsilon_{f_t} \quad (5)$$

donde $\epsilon_{f_t} \sim N(0,1)$ y el logaritmo de la varianza del factor (σ_{f_t}) sigue un paseo aleatorio como el descrito a continuación

$$\log \sigma_{f_t} = \log \sigma_{f_{t-1}} + v_{f_t} \quad ; \quad v_{f_t} \sim N(0, \omega_{f_t}) \quad (6)$$

La incorporación de volatilidad estocástica en modelos de factores dinámicos con mezcla de frecuencias fue propuesta por Marcellino, Porqueddu y Venditti (2016) y, posteriormente, por Pacce y Pérez Quirós (2019) en el marco del modelo Euro-STING (Camacho y Pérez Quirós, 2010). En el presente trabajo se sigue la segunda línea metodológica, lo que implica que es necesario reescribir la especificación descrita en las ecuaciones (1), (3), (5) y (6) utilizando la representación de vectores apilados¹⁰. Esta especificación permite la estimación bayesiana del modelo haciendo más sencilla la estimación cuando se incorpora la volatilidad estocástica en el modelo. El anejo 2 presenta una descripción detallada de la estimación bayesiana del modelo, la cual se realiza a través del empleo de un algoritmo Metropolis-Hastings a partir de un muestreo tipo Gibbs¹¹.

¹⁰ Para la estimación bayesiana de modelos de factores dinámicos con mezcla de frecuencias se ha seguido la aproximación de vectores apilados propuesta por Koopman y Pacce (2015). En el anejo 1 se explica la forma del modelo especificado en las ecuaciones (1), (3), (5) y (6) siguiendo la mencionada aproximación.

¹¹ Para la estimación del componente estocástico de la volatilidad del factor, se sigue la metodología descrita en Kim, Shepard y Chib (1998).

Es importante tener en cuenta que la solución propuesta tiene como objetivo principal, no solo mejorar las predicciones en el período posterior al de la pandemia, sino también que el modelo no se comporte peor respecto a las predicciones realizadas en el período previo a la pandemia. Sin embargo, no es un objetivo identificar la especificación de un modelo que permita obtener resultados satisfactorios en situaciones de alta incertidumbre, como la que se experimentó a lo largo de 2020. Para alcanzar este propósito, sería necesario adoptar un enfoque diferente o el uso de otras herramientas econométricas que consideren información económica con una frecuencia más elevada.

Alternativamente a la inclusión de volatilidad estocástica en el factor común, se puede considerar la asignación de observaciones faltantes durante el período más crítico de la pandemia del COVID-19, período que mostró los mayores cambios en la dinámica de las variables. En otras palabras, ello significa dejar incompleta la información económica procedente de los indicadores durante ese período. Esta alternativa fue considerada dentro de la evaluación de las distintas alternativas de modelización, si bien se decidió descartarla debido a que presentaba resultados menos satisfactorios que los que provenían de la modelización que incorpora la volatilidad estocástica en el modelo. En el anejo 3 se encuentra una explicación detallada de esta opción y un análisis comparativo de los resultados obtenidos a partir de ella.

Para analizar la conveniencia de la introducción de volatilidad estocástica en el modelo, se presenta un conjunto seleccionado de pruebas relevantes con el objetivo de evaluar la capacidad predictiva de las alternativas de modelización descritas. En concreto, el modelo *Anterior* y el modelo con volatilidad estocástica —en adelante, «modelo SV»¹²—. Partiendo de las variables incluidas en el modelo de Arencibia Pareja, Gómez Loscos, De Luis López y Pérez Quirós (2020) (véase cuadro 1)¹³, el análisis se centra en la comparación de la capacidad predictiva del modelo *Anterior* respecto a la especificación que introduce volatilidad estocástica en el componente común del modelo. No obstante, para aislar las potenciales mejoras provenientes de la incorporación de la volatilidad estocástica de los cambios en las relaciones temporales de las variables descritas en el epígrafe anterior, se incorpora el PMI compuesto, el ESI sin consumidores y el IPI de construcción con una relación contemporánea con el factor común en ambas especificaciones.

El cuadro 3 presenta los errores de previsión de ambos modelos para los diferentes períodos. En particular, se considera la raíz de los errores cuadráticos medios y los errores

12 Cabe señalar que se ha diseñado y evaluado la capacidad predictiva de un conjunto más amplio de especificaciones que el mostrado en el presente epígrafe. No obstante, estos resultados no se exponen de forma detallada en este documento debido a la baja capacidad predictiva que presentan. En particular, no se describen los resultados obtenidos con dos especificaciones alternativas de modelización: la primera incorpora volatilidad estocástica tanto en el factor como en las variables, y la segunda la incluye únicamente en las variables del modelo, mientras la volatilidad asociada al factor se mantiene constante. Los resultados de estos ejercicios están disponibles bajo petición a los autores.

13 Entre marzo de 2020 y marzo de 2022 se facilitó en España el uso de los Expedientes de Regulación de Empleo Temporal (ERTE), que posibilitaban la suspensión temporal de un contrato de trabajo sin necesidad de recurrir al despido. La serie bruta de afiliados a la Seguridad Social no excluye aquellos contratos de trabajo sujetos a ERTE durante el período mencionado. Con la intención de que la serie refleje de la manera más fidedigna posible el nivel de actividad económica en España y, por tanto, tenga un mayor poder explicativo sobre el PIB, a partir de febrero de 2020 se excluye de la serie de afiliados a aquellos trabajadores acogidos a un régimen de ERTE.

Cuadro 3

ERRORES DE PREDICCIÓN EN LOS PERÍODOS PRE- Y POS-COVID

		Pre-COVID (I TR 2015 - IV TR 2019)		Pos-COVID (I TR 2021 - II TR 2023)		Pos-COVID (I TR 2021 - II TR 2023) excluyendo II TR 2021	
		PIB (estimación <i>flash</i>)	PIB (2.ª estimación)	PIB (estimación <i>flash</i>)	PIB (2.ª estimación)	PIB (estimación <i>flash</i>)	PIB (2.ª estimación)
Error medio absoluto	Modelo <i>Anterior</i>	0,10	0,12	0,67	0,67	0,61	0,67
	Modelo SV	0,08	0,10	0,25	0,46	0,25	0,35
Raíz del error cuadrático medio	Modelo <i>Anterior</i>	0,11	0,14	0,85	0,84	0,82	0,86
	Modelo SV	0,11	0,13	0,30	0,60	0,30	0,41

FUENTES: Elaboración propia.

NOTA: Los errores cuadráticos medios y el error medio absoluto que se muestran están relacionados con las previsiones realizadas en el tercer mes del trimestre objetivo. El modelo *Anterior* se refiere al modelo descrito en el epígrafe 2, pero incluyendo el ISE sin consumidores, el PMI compuesto y el IPI de construcción con una relación contemporánea (y no adelantada) respecto al factor común. El modelo SV se refiere al modelo que incluye volatilidad estocástica en el factor. Para el cómputo del error cuadrático medio y de la raíz del error cuadrático medio de la séptima y octava columna se excluye el error cometido en el segundo trimestre de 2021.

medios absolutos, calculados a partir de un ejercicio en tiempo real como el descrito en el epígrafe 3.1. Es decir, se tienen en cuenta los errores cometidos con las estimaciones realizadas en el tercer mes de cada uno de los trimestres, momento en el cual se cuenta, en función de la variable, con uno o dos meses de información del trimestre en curso. De acuerdo con los resultados obtenidos, el modelo SV presenta una mejor capacidad predictiva que el modelo *Anterior*. Por un lado, en el caso del período posterior al COVID-19 (2021-2023), el modelo SV permite obtener una mejora significativa en la raíz de errores medios al cuadrado y también en el valor del error medio absoluto, tanto en el caso de la segunda estimación del PIB como, especialmente, en el de la estimación *flash*. Por otro lado, durante el período previo a la pandemia (2015-2019) se observa que el modelo SV reduce marginalmente los errores de previsión del modelo *Anterior*. Finalmente, cabe destacar que, cuando se omite del cómputo el error cometido en el segundo trimestre de 2021 —trimestre en el cual el cambio entre la estimación *flash* y la segunda estimación del PIB alcanzó casi 2 pp (véase gráfico 4)—, la mejora observada en cuanto a la capacidad predictiva del modelo SV se incrementa notablemente para el caso de los errores cometidos respecto a la segunda estimación del PIB.

3.3 Modificación del conjunto de variables del modelo

Con el objetivo de analizar la posibilidad de lograr mejoras adicionales en la previsión del PIB, se considera la alternativa de que algunos de los indicadores incluidos hasta ahora en el modelo hayan perdido poder explicativo, al mismo tiempo que otros, que no se encontraban en ese conjunto de variables, puedan mejorar la previsión trimestral del PIB. Cabe señalar que, dada la colinealidad de los indicadores macroeconómicos más habituales, la inclusión tentativa de variables adicionales no tiene por qué mejorar la capacidad predictiva del modelo. Consecuentemente, es recomendable el empleo de un procedimiento de selección estadístico que permita identificar aquellas variables que son relevantes de acuerdo con determinados criterios estadísticos. En el presente documento, se sigue la metodología

ERRORES DE PREDICCIÓN PARA EL MODELO SV PARTIENDO DEL MODELO BASE Y AÑADIENDO INDICADORES

		Pre-COVID (I TR 2015 - IV TR 2019)		Pos-COVID (I TR 2021 - II TR 2023)	
		PIB (estimación <i>flash</i>)	PIB (2.ª estimación)	PIB (estimación <i>flash</i>)	PIB (2.ª estimación)
Error medio absoluto	(1) Modelo base	0,10	0,12	0,43	0,65
	(2) Modelo base + ISE sin consumidores	0,10	0,11	0,42	0,63
	(3) Modelo base + ISE sin consumidores + Ventas de grandes empresas	0,09	0,10	0,40	0,68
	(4) Modelo base + ISE sin consumidores + Ventas de grandes empresas + IPI de construcción	0,08	0,10	0,24	0,46
Raíz del error cuadrático medio	(1) Modelo base	0,14	0,16	0,57	1,01
	(2) Modelo base + ISE sin consumidores	0,13	0,15	0,54	0,98
	(3) Modelo base + ISE sin consumidores + Ventas de grandes empresas	0,11	0,13	0,49	0,92
	(4) Modelo base + ISE sin consumidores + Ventas de grandes empresas + IPI de construcción	0,11	0,13	0,29	0,60

FUENTES: Elaboración propia.

NOTA: Se muestran los errores cuadráticos medios y el error medio absoluto relacionado con las previsiones realizadas en el tercer mes del trimestre objetivo. Todas las especificaciones se refieren al modelo que incluye volatilidad estocástica en el factor. El modelo base incluye las siguientes variables: PIB, afiliados a la Seguridad Social, consumo eléctrico, PMI compuesto, IPI no energético, exportaciones reales de bienes e importaciones reales de bienes. En el resto de los modelos se especifica la variable o variables añadidas al modelo base para la estimación.

propuesta por Camacho y Pérez-Quirós (2010)¹⁴. En concreto, se parte de un modelo base que incluye un conjunto parsimonioso de variables representativo de la evolución de la actividad y se estudia si la incorporación de variables adicionales logra mejorar o no la capacidad predictiva del modelo en el período posterior al COVID-19, sin deteriorar la correspondiente al período previo a la pandemia.

Así, el conjunto de variables del modelo base se selecciona con la intención de contar con indicadores de actividad (PIB y consumo eléctrico), de oferta (IPI no energético), de empleo (afiliados a la Seguridad Social), de demanda exterior e interior (exportaciones e importaciones reales de bienes), y al menos uno de opinión, con una publicación próxima a la finalización del mes de referencia (PMI compuesto). Por su parte, se evalúa la incorporación a este modelo base, de uno en uno, del resto de los indicadores que se encontraban incluidos en la versión anterior del modelo: el ISE sin consumidores, las Ventas de grandes empresas y el IPI de construcción. En cada etapa se analiza la capacidad predictiva del

¹⁴ Alvarez, Camacho y Pérez-Quirós (2012) examinan los pros y los contras empíricos de la previsión con modelos de factores grandes frente a modelos de factores pequeños, y obtienen que, cuanto mayor es el número de series temporales, mayor es la correlación entre ellas y, por tanto, mayor es la correlación de la parte idiosincrásica, siendo esta correlación la que podría sesgar los resultados del factor común estimado. Además, Bai y Ng (2008) han demostrado la importancia de contar con especificaciones parsimoniosas para mejorar la capacidad de previsión de los modelos factoriales, incluso cuando se mantiene una correlación cruzada igual a cero en la parte idiosincrásica. Por último, Boivin y Ng (2006) muestran que las ventajas asintóticas de los modelos factoriales a gran escala distan mucho de mantenerse en las aplicaciones empíricas.

modelo, en tiempo real, durante el período previo y posterior al COVID-19, teniendo en cuenta los valores observados tanto para la raíz del error cuadrático medio como para el error medio absoluto. De esta forma, el criterio de decisión para la incorporación o no de una determinada variable es que su inclusión permita reducir el error de las predicciones del modelo, esto es, que ayude a mejorar su capacidad predictiva.

En el cuadro 4 se presentan el error medio absoluto y la raíz del error cuadrático medio de cada modelo estimado, que incorporan en todos los casos tanto la relación contemporánea de todas las variables con el factor común como la volatilidad estocástica en la dinámica del componente común estimado. Así, se observa que, en general, conforme se van añadiendo las variables mencionadas, los errores de predicción se reducen ligeramente tanto en el período anterior como en el período posterior al COVID-19, lo que es indicativo del poder explicativo de cada una de las variables incorporadas en el modelo. Por su parte, la inclusión del IPI de construcción en el modelo es la que logra una reducción más significativa en los errores de previsión del modelo en el período posterior a la pandemia (I TR 2021- II TR 2023).

Cabe destacar que el último modelo mencionado muestra errores de previsión marginalmente menores que los que se exhiben en el cuadro 3, los cuales provienen del mismo modelo, pero incluyendo además el crédito a sociedades no financieras. Adicionalmente, esta variable presenta un signo de factor de carga asociado que dificulta su interpretación económica, ya que, como se menciona en el epígrafe 2.2, a partir de la inclusión de la pandemia en el período de análisis dicho signo pasó a ser negativo y, por tanto, sugeriría que incrementos en el crédito se relacionan con caídas de la actividad¹⁵. Por estos motivos, se decide excluir la variable crédito a sociedades no financieras del modelo final.

Además del ejercicio descrito, se realizan dos ejercicios adicionales. Por un lado, se estudia la incorporación al modelo del Indicador de Actividad del Sector Servicios (IASS), variable que, al medir la actividad en los servicios —sector cuyo peso en el PIB es significativo—, podría ser de utilidad para mejorar la capacidad predictiva del modelo. Así, sobre la última especificación que se muestra en el cuadro 4 (especificación 4), se añade el IASS. Si bien los resultados apuntan a que este indicador podría lograr una ligera mejora en la capacidad predictiva del modelo, el factor de carga asociado a él presenta el signo contrario al esperado en el período pos-COVID. Consecuentemente, se descarta la inclusión de este indicador como potencial variable adicional del modelo. Por otro lado, se analiza la posibilidad de sustituir el indicador de ventas de grandes empresas por el del Índice de Comercio al por Menor (ICM). Este ejercicio se realiza también sobre la especificación (4) del cuadro 4. En este caso, los resultados de estas pruebas indican que la especificación que incorpora el ICM empeora de forma notable la capacidad predictiva del modelo.

15 El cambio de signo en el factor de carga puede deberse a que, en los momentos más críticos de la pandemia —cuando la caída de la actividad fue más intensa—, el crédito en España se mantuvo dinámico como resultado de las políticas económicas centradas en el apoyo a las empresas y familias implementadas por las autoridades.

Cuadro 5

INDICADORES UTILIZADOS EN EL MODELO SPAIN-STING REVISADO

Indicador	Tipo de indicador	Fuente	Periodicidad	Relación	Fecha de inicio	Retraso en su publicación
Tasa de variación del PIB	Actividad	INE	Trimestral	Contemporánea	Marzo 1990	+ 30 días
ISE sin consumidores	Basado en encuesta	Comisión Europea	Mensual	Contemporánea	Enero 1990	0 días
PMI compuesto	Basado en encuesta	IHS Markit	Mensual	Contemporánea	Agosto 1990	+ 5 días
Consumo eléctrico	Actividad	Red Eléctrica de España	Mensual	Contemporánea	Febrero 1990	+ 1 día
Afiliados a la Seguridad Social	Actividad	Seguridad Social	Mensual	Contemporánea	Enero 1990	+ 3 días
Ventas de grandes empresas	Actividad	Agencia Tributaria	Mensual	Contemporánea	Febrero 1996	+ 10 días
IPI no energético	Actividad	INE	Mensual	Contemporánea	Febrero 1992	+ 36 días
IPI de construcción	Actividad	INE	Mensual	Contemporánea	Febrero 1992	+ 36 días
Importaciones reales de bienes	Actividad	Departamento de Aduanas y MINECO	Mensual	Contemporánea	Febrero 1991	+ 50 días
Exportaciones reales de bienes	Actividad	Departamento de Aduanas y MINECO	Mensual	Contemporánea	Febrero 1991	+ 50 días

FUENTES: Elaboración propia.

Por lo tanto, una vez analizado el conjunto de pruebas relevante, la especificación que muestra los menores errores de predicción estimados como errores medios absolutos o como raíz de los errores cuadráticos medios, y que además mantiene un signo coherente con la teoría económica en los factores de carga asociados a cada una de las variables, es la que incluye todos los indicadores con una relación contemporánea con el factor común y contiene las siguientes variables: PIB, afiliados a la Seguridad Social, ventas de grandes empresas, consumo eléctrico, IPI no energético, PMI compuesto, exportaciones reales de bienes, importaciones reales de bienes, ISE sin consumidores e IPI de construcción (véase cuadro 5).

3.4 Evaluación de la capacidad predictiva del modelo *Revisado* frente al *Anterior*

Una vez explicados los tres cambios realizados en el modelo y analizada la mejora de la capacidad predictiva de cada uno de ellos, resulta fundamental comparar la capacidad predictiva del modelo *Revisado* —el que incorpora de forma simultánea los tres cambios descritos— con la capacidad predictiva del modelo *Anterior*¹⁶.

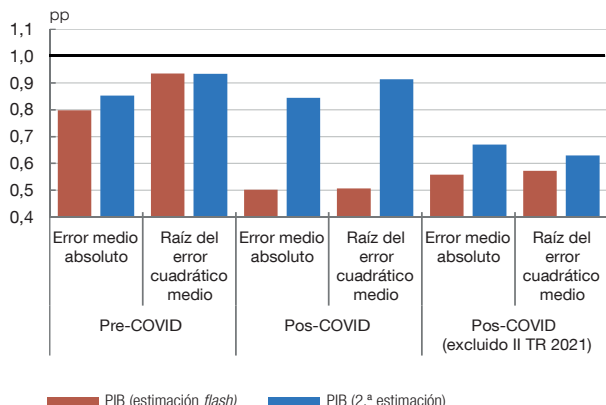
En el gráfico 5.1 se muestran los errores de previsión, en tiempo real, del modelo *Revisado* en términos relativos a los errores correspondientes al modelo *Anterior*. En los casos en los que el valor de la relación es menor que uno, el modelo *Revisado* muestra una mejor capacidad predictiva frente al modelo *Anterior*, al presentar errores de previsión de menor magnitud, y lo contrario ocurre cuando el valor es mayor que la unidad. Más concretamente, se analizan, por un lado, los errores cometidos entre el primer trimestre de 2015 y el último de 2019 (período pre-COVID) y, por otro lado, los errores cometidos entre el primer trimestre

¹⁶ El modelo descrito en el epígrafe 2, el cual contiene variables relacionadas de forma adelantada respecto al factor común, no permite que la varianza del factor común sufra cambios en el tiempo, e incluye como indicador el crédito a sociedades no financieras.

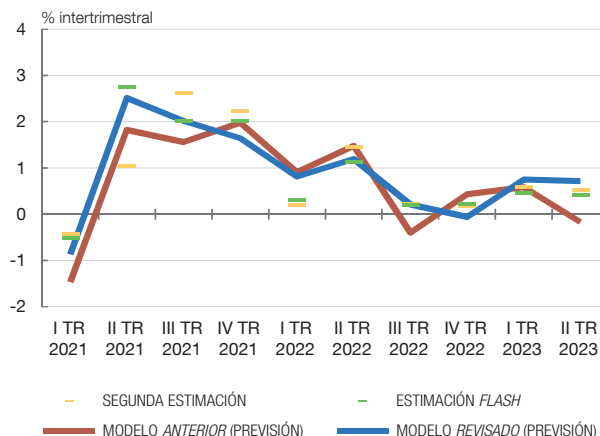
Gráfico 5

MEJORAS EN LA PREVISIÓN DEL MODELO REVISADO FRENTE AL MODELO ANTERIOR

1 ERRORES DE PREDICCIÓN RELATIVOS. MODELO REVISADO EN RELACIÓN CON EL MODELO ANTERIOR (PREVISIÓN EN TIEMPO REAL)



2 PIB: OBSERVADO Y PREVISIONES. PERÍODO 2021-2023



FUENTE: Elaboración propia.

NOTA: Se representan los errores de predicción del modelo *Revisado* en relación con los errores del modelo descrito en el epígrafe 2. Se tienen en consideración los errores cometidos con información hasta mediados del tercer mes de cada trimestre. El período «pre-COVID» se refiere a los trimestres comprendidos entre el primero de 2015 y el último de 2019. «Pos-COVID» comprende desde el primer trimestre de 2021 hasta el segundo de 2023. «Pos-COVID (excluido II TR 2021)» se refiere a excluir del cómputo de los errores del período «Pos-COVID» al error cometido en el segundo trimestre de 2021, dada la intensa revisión que se hizo en ese trimestre entre la estimación *flash* y la segunda estimación del PIB.

de 2021 y el segundo de 2023 (período pos-COVID), tanto cuando la comparación es frente a la estimación *flash* del PIB como cuando es frente a la segunda estimación.

El primer resultado que hay que destacar es que, en todos los casos analizados, los errores de previsión correspondientes al modelo *Revisado* son inferiores a los cometidos por el modelo *Anterior*, lo que confirma las ganancias en términos de capacidad predictiva de la nueva especificación del modelo. Si se analizan los resultados para los diferentes períodos, se observa que la mejora lograda por el modelo *Revisado* para el período pre-COVID (de entre un 10 % y un 20 %, dependiendo de si se analiza el error medio absoluto o el error cuadrático medio) no solo cumple con el objetivo inicial de no empeorar la capacidad predictiva en dicho período, sino que incluso logra una cierta mejora de esta. En cuanto al período pos-COVID, se observa una diferencia significativa en los resultados cuando la intención es predecir la estimación *flash* del PIB respecto a los obtenidos cuando se quiere predecir la segunda estimación del PIB. En concreto, los errores asociados al modelo *Revisado* en el primer caso son un 50 % inferiores a los cometidos por el modelo *Anterior*, mientras que para la segunda estimación del PIB las mejoras logradas serían cercanas al 10 %. No obstante, como se aprecia en el gráfico 5.2, donde se muestra la comparación de las previsiones de los modelos respecto a la segunda estimación y la estimación *flash* del PIB para cada uno de los trimestres en el período posterior a la pandemia, este resultado está muy influido por la inclusión en el período de análisis del segundo trimestre de 2021, en el cual se observó una revisión de más de 1,5 pp en la tasa de crecimiento intertrimestral del PIB entre la primera y la segunda estimación. Excluyendo ese trimestre de los cálculos de los errores de previsión, la reducción lograda

Cuadro 6

REGRESIONES PARA LA ESTIMACIÓN DEL PIB EN LOS PERÍODOS PRE-COVID (I TR 2015 - IV TR 2019) Y POS-COVID (I TR 2021 - II TR 2023)

		PIB (estimación <i>flash</i>)		PIB (2.ª estimación)	
		Modelo <i>Anterior</i>	Modelo <i>Revisado</i>	Modelo <i>Anterior</i>	Modelo <i>Revisado</i>
Pre-COVID (I TR 2015 - IV TR 2019)	Constante	0,22	0,13	0,33 **	0,27 **
		[0,13]	[0,12]	[0,12]	[0,11]
	Previsión	-1,64	-1,07	(-2,99)	-2,42
		0,66 *	0,76	0,51 **	0,59 **
		[0,18]	[0,17]	[0,16]	[0,15]
		(-1,88)	(-1,43)	(-2,99)	(-2,67)
	R ²	0,43	0,54	0,36	0,44
R ² ajustado	0,40	0,51	0,32	0,41	
Número de observaciones	20	20	20	20	
Pos-COVID (I TR 2021 - II TR 2023)	Constante	-0,16	0,06	-0,12	0,18
		[0,25]	[0,13]	[0,30]	[0,28]
	Previsión	(-0,66)	(0,45)	(-0,40)	(0,65)
		0,93	0,92	0,92	0,82
		[0,19]	[0,10]	[0,24]	[0,22]
		(-0,40)	(-0,76)	(-0,35)	(-0,81)
	R ²	0,76	0,91	0,65	0,64
R ² ajustado	0,73	0,90	0,61	0,59	
Número de observaciones	10	10	10	10	

FUENTES: Elaboración propia.

NOTA: El período «pre-COVID» hace referencia a los errores obtenidos entre el primer trimestre de 2015 y el cuarto de 2019. El período «pos-COVID» se refiere a los trimestres comprendidos entre el primer trimestre de 2021 y el primero de 2023. En todos los casos se consideran las previsiones obtenidas a mediados del tercer mes de cada uno de los trimestres del período. Entre corchetes está representado el error estándar, y entre paréntesis se recoge el t-estadístico para cada una de las hipótesis nulas. En el caso de la constante, *, ** y *** indican que es significativamente diferente de cero para un nivel de confianza del 90 %, 95 % y 99 %, respectivamente. Por su parte, *, ** y *** indican que el coeficiente asociado al regresor es significativamente diferente de la unidad para un nivel de confianza del 90 %, 95 % y 99 %, respectivamente. El «modelo *Anterior*» se refiere al modelo descrito en el epígrafe 2, pero incluyendo el ISE sin consumidores, el PMI compuesto y el IPI de construcción con una relación contemporánea (y no adelantada) respecto al factor común. El «modelo *Revisado*» se refiere al modelo que incluye volatilidad estocástica en el factor, excluye la variable de crédito y relaciona todas las variables con el factor de forma contemporánea.

por el modelo *Revisado* en el error de previsión de la segunda estimación del PIB, respecto al modelo *Anterior*, es de entre un 25 % y un 35 %.

Complementariamente al estudio de los errores de predicción, es posible analizar los sesgos potenciales en las previsiones obtenidas por los dos modelos considerados mediante regresiones que tienen en cuenta como variable dependiente la estimación *flash* o la segunda estimación del PIB y, como variable explicativa, las previsiones de los diferentes modelos (véase cuadro 6). Nuevamente, en el análisis se distingue entre los resultados para el período pre-COVID (panel superior) y los resultados para el período pos-COVID (panel inferior).

En el período previo a la pandemia, cuando la variable dependiente es la estimación *flash* del PIB, se observa que no existe un sesgo sistemático en las previsiones, ya que no se puede rechazar la hipótesis de que la constante de los modelos analizados sea distinta de cero a ningún nivel de significatividad estadística. Esto no sucede cuando la variable

dependiente es la segunda estimación del PIB, aunque el sesgo aparente es reducido. Por otro lado, el valor positivo y elevado de los coeficientes asociados a la variable Previsión indica que las previsiones que surgen tanto del modelo *Anterior* como del que incorpora las modificaciones descritas en los epígrafes anteriores son buenos predictores. Además, el modelo *Revisado* parece comportarse relativamente mejor, en tanto que dicho coeficiente no es estadísticamente diferente de la unidad en el caso en el que la variable dependiente es la estimación *flash*. Por último, el R-cuadrado ajustado es más elevado en el modelo *Revisado*. En las estimaciones que tienen en cuenta el período pos-COVID¹⁷, independientemente de si se analizan los resultados para la estimación *flash* o para la segunda estimación del PIB, se identifican los siguientes hechos estilizados. En primer lugar, no se observa un sesgo sistemático en las previsiones, ya que la constante no es significativamente distinta de cero, en todas las especificaciones analizadas. En segundo lugar, no puede rechazarse que el coeficiente asociado a la previsión, tanto del modelo *Anterior* como del modelo *Revisado*, sea distinto a la unidad desde un punto de vista estadístico, lo que indica nuevamente el buen ajuste de las previsiones que surgen de los modelos.

Por último, el R-cuadrado ajustado del modelo con volatilidad estocástica presenta el mayor valor en el caso en el que la variable dependiente es la estimación *flash*, si bien en el caso en que la variable dependiente es la segunda estimación del PIB es el modelo *Anterior* el que mejor ajuste muestra desde el punto de vista del estadístico mencionado, aunque la diferencia es relativamente pequeña.

Por lo tanto, los resultados de las pruebas aportadas permiten concluir que las modificaciones introducidas en el modelo consiguen mejorar las previsiones de la tasa intertrimestral del PIB desde el primer trimestre de 2021, al mismo tiempo que se mejoran marginalmente las del período previo a la pandemia. Asimismo, el modelo *Revisado* no parece mostrar sesgos en las predicciones, a la vez que mejora el ajuste a los datos de acuerdo con los R-cuadrado ajustados.

¹⁷ Es necesario tener en cuenta que el período pos-COVID incorpora un número reducido de observaciones y, consecuentemente, estos resultados deben ser interpretados con cautela.

4 Consideraciones finales

Las previsiones a corto plazo sobre el curso futuro de la economía desempeñan un papel crucial en la toma de decisiones de los bancos centrales y de otras instituciones nacionales e internacionales. Las disrupciones vinculadas a la pandemia acarrearón un aumento sin precedentes de la volatilidad de los indicadores económicos, lo que redundó en un deterioro de la capacidad predictiva de los modelos de previsión a corto plazo. En el caso de Spain-STING, el cambio en la dinámica de las variables y el aumento de la volatilidad ha afectado a la relación de largo plazo entre los indicadores y el componente común estimado por el modelo, lo cual ha implicado una reducción significativa de su capacidad predictiva en el período pos-COVID (2021-2023).

En este trabajo, se han presentado tres cambios fundamentales en la especificación del modelo Spain-STING que permiten mejorar su capacidad predictiva tras la pandemia. En concreto: i) se considera que todas las variables empleadas en la predicción evolucionan de manera contemporánea en el tiempo con el componente común identificado —en lugar de incluir algunas de ellas de forma adelantada—; ii) se incorpora volatilidad estocástica en el componente común del modelo, y iii) se reevalúa el conjunto de variables incluidas en el modelo.

En términos cuantitativos, la combinación de las tres modificaciones consigue que los errores de predicción se reduzcan en el período pos-COVID, respecto al modelo *Anterior*, entre un 10 % y un 50 %, en función de si se evalúa la predicción de la estimación *flash* o la de la segunda estimación del PIB. Así, la consideración simultánea de las modificaciones propuestas en este documento resulta en una especificación del modelo que corrige de forma notable el deterioro de la capacidad predictiva observado entre 2021 y 2023. Aun así, estos resultados deben ser interpretados con cautela. Por un lado, el período pos-COVID cuenta con un número reducido de observaciones y, por otro lado, existe incertidumbre con respecto a si los cambios observados en la dinámica de las variables son de naturaleza transitoria, debidos a la pandemia, o de una persistencia más duradera. En todo caso, los resultados obtenidos muestran también una cierta mejora (del entorno del 10 %) en la capacidad predictiva del modelo en el período previo a la pandemia, cuando la volatilidad de las variables era más reducida.

Bibliografía

- Alvarez, Rocio, Maximo Camacho y Gabriel Pérez-Quirós. (2012). "Finite Sample Performance of Small Versus Large Scale Dynamic Factor Models". CEPR Discussion Paper, DP8867. <https://doi.org/10.2139/ssrn.2002663>
- Arencibia Pareja, Ana, Ana Gómez Loscos, Mercedes de Luis López y Gabriel Pérez Quirós. (2020). "A short term forecasting model for the Spanish GDP and its demand components". *Revista Economía*, 43(85), pp. 1-30. <https://doi.org/10.18800/economia.202001.001>
- Bai, Jushan, y Serena Ng. (2008). "Forecasting economic time series using targeted predictors". *Journal of Econometrics*, 146, pp. 304-317. <https://doi.org/10.1016/j.jeconom.2008.08.010>
- Boivin, Jean, y Serena Ng. (2006). "Are more data always better for factor analysis?". *Journal of Econometrics*, 132, pp. 169-194. <https://doi.org/10.1016/j.jeconom.2005.01.027>
- Camacho, Maximo, y Gabriel Pérez-Quirós. (2010). "Introducing the euro-STING: Short-Term Indicator of euro area Growth". *Journal of Applied Econometrics*, 25(4), pp. 663-694. <https://doi.org/10.1002/jae.1174>
- Camacho, Maximo, y Gabriel Pérez-Quirós. (2011). "Spain-STING: Spain Short-Term Indicator of Growth". *The Manchester School*, 79(1), pp. 594-616. <https://doi.org/10.1111/j.1467-9957.2010.02212.x>
- Carter, C. K., y R. Kohn. (1994). "On gibbs sampling for state space models". *Biometrika*, 81(3), pp. 541-553. <https://doi.org/10.2307/2337125>
- Comisión Europea. (2006). "The joint harmonized EU programme of business and consumer surveys". Special Report, 5/2006. https://ec.europa.eu/economy_finance/db_indicators/surveys/documents/studies/ee_bcs_2006_05_en.pdf
- Durbin, J., y S. J. Koopman. (2002). "A simple and efficient simulation smoother for state space time series analysis". *Biometrika*, 89(3), pp. 603-616. <https://doi.org/10.1093/biomet/89.3.603>
- Hamilton, James D. (1994). "State-space models". En Robert F. Engle y Daniel L. McFadden (eds.), *Handbook of Econometrics*, 4, pp. 3039-3080. [https://doi.org/10.1016/S1573-4412\(05\)80019-4](https://doi.org/10.1016/S1573-4412(05)80019-4)
- Kim, Chang-Jin, y Charles R. Nelson. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. The MIT Press. <https://doi.org/10.7551/mitpress/6444.001.0001>
- Kim, Sangjoon, Neil Shephard y Siddhartha Chib. (1998). «Stochastic volatility: likelihood inference and comparison with ARCH models». *The Review of Economic Studies*, 65(3), pp. 361-393. <https://doi.org/10.1111/1467-937X.00050>
- Koopman, Siem Jan, y Matías J. Pacce. (2016). *Short-term forecasting of business cycle turning points: a mixed-frequency Markov-switching dynamic factor analysis*. 9th ECB Workshop on Forecasting Techniques: Forecast Uncertainty and Macroeconomic Indicators, 3 y 4 de junio de 2016. Banco Central Europeo. https://www.ecb.europa.eu/pub/conferences/shared/pdf/20160603_forecasting/Paper_6_Pacce.pdf
- Marcellino, Massimiliano, Mario Porqueddu y Fabrizio Venditti. (2016). "Short-term GDP forecasting with a mixed-frequency dynamic factor model with Stochastic Volatility". *Journal of Business & Economic Statistics*, 34(1), pp. 118-127. <https://doi.org/10.1080/07350015.2015.1006773>

- Mariano, Roberto, y Yasutomo Murasawa. (2003). "A new coincident index of business cycles based on monthly and quarterly series". *Journal of Applied Econometrics*, 18 (4), pp. 427-443. <https://doi.org/10.1002/jae.695>
- Maroz, Danila, James Stock y Mark Watson. (2021). "Comovement of Economic Activity During the Covid Recession". Manuscript. https://www.princeton.edu/~mwatson/papers/Covid_Factor_20211215.pdf
- Pacce, Matías José, y Gabriel Pérez Quirós. (2019). "Predicción en tiempo real del PIB en el área del euro: recientes mejoras en el modelo Euro-STING". *Boletín Económico - Banco de España*, 1/2029, Notas Económicas. <https://repositorio.bde.es/handle/123456789/8443>
- Schorfheide, Frank, y Son Dongho. (2021). "Real-Time Forecasting with a (Standard) Mixed-Frequency VAR during a Pandemic". Working Paper, 29535, National Bureau of Economic Research. <https://doi.org/10.3386/w29535>

Anejo 1. Aproximación de vectores apilados para la representación del modelo

Siguiendo la aproximación de vectores apilados, las ecuaciones (1) y (3) pueden escribirse como:

$$\begin{pmatrix} X_{\tau}^1 \\ Z_{\tau,3}^2 \\ Z_{\tau,2}^2 \\ Z_{\tau,1}^2 \end{pmatrix} = \begin{bmatrix} 1/3\beta_1 & 2/3\beta_1 & \beta_1 & 2/3\beta_1 & 1/3\beta_1 \\ \beta_2 & 0 & 0 & 0 & 0 \\ 0 & \beta_2 & 0 & 0 & 0 \\ 0 & 0 & \beta_2 & 0 & 0 \end{bmatrix} \begin{bmatrix} f_{\tau,3} \\ f_{\tau,2} \\ f_{\tau,1} \\ f_{\tau-1,1} \\ f_{\tau-1,2} \end{bmatrix} + \begin{bmatrix} u_{\tau}^1 \\ u_{\tau,3}^1 \\ u_{\tau,2}^1 \\ u_{\tau,1}^1 \end{bmatrix} \quad (7)$$

$$\phi_1(L) u_{\tau}^1 = \varepsilon_{\tau}^1 \quad (8)$$

$$\phi_2(L) u_{\tau,3}^2 = \varepsilon_{\tau,3}^2 \quad (9)$$

donde el subíndice (τ, k) con $k = 1, 2, 3$ se refiere al trimestre $\tau = 1, \dots, T/3$ y el mes 1, 2, 3 dentro de cada trimestre [por ejemplo, si τ es el primer trimestre de 2023, entonces $(\tau, 1)$ es enero de 2023]. Nótese que esta forma de representación permite que el error asociado a la variable trimestral (u_{τ}^1) tenga frecuencia trimestral y el operador L actúe sobre dicha frecuencia, mientras que el operador L sobre $u_{\tau,3}^2$ sigue afectando a los errores sobre una frecuencia mensual como en la ecuación (3). Este pequeño cambio es de vital importancia para la estimación del modelo bajo una aproximación bayesiana. En particular, la dinámica de la variable trimestral puede representarse como:

$$\phi_1(L) X_{\tau}^1 = \phi_1(L) \beta_1 \left(\frac{1}{3} f_{\tau,3} + \frac{2}{3} f_{\tau,2} + f_{\tau,1} + \frac{2}{3} f_{\tau-1,3} + \frac{1}{3} f_{\tau-1,2} \right) + \varepsilon_{\tau}^1 \quad (10)$$

donde el error de la ecuación es de ruido blanco y, por tanto, es posible utilizar las especificaciones bayesianas habituales para la estimación de los parámetros β_1 y σ_1 ¹⁸. Para un mayor detalle de las ganancias asociadas a esta representación en la estimación de modelos de factores dinámicos con mezcla de frecuencias, véase Koopman y Pacce (2015).

En el caso de la ecuación (5), la estimación de los parámetros autorregresivos asociados a la dinámica del error del factor común cuando se incorpora volatilidad estocástica puede realizarse simplemente despejando ε_{τ} , de forma que

$$\frac{\phi_1(L) f_{\tau}}{\sigma_{f_{\tau}}} = \varepsilon_{\tau} \quad (11)$$

donde ε_{τ} es un ruido blanco.

¹⁸ Nótese que, si se parte de la representación de la ecuación (1), la premultiplicación de dicha ecuación por el polinomio de retardos correspondiente da como resultado:

$$\phi_1(L) X_{\tau}^1 = \phi_1(L) \beta_1 (1/3f_{\tau} + 2/3f_{\tau-1} + f_{\tau} + 2/3f_{\tau-2} + 1/3f_{\tau-3}) + (1/3\varepsilon_{\tau}^1 + 2/3\varepsilon_{\tau-1}^1 + \varepsilon_{\tau-2}^1 + 2/3\varepsilon_{\tau-3}^1 + 1/3\varepsilon_{\tau-4}^1)$$

siendo la dinámica del error asociado a $\phi_1(L) X_{\tau}^1$ de forma MA(4) y, por tanto, difícil de estimar desde un punto de vista bayesiano.

Anejo 2. Estimación bayesiana

Para la estimación del modelo se emplea un algoritmo Metropolis-Hastings a partir de un muestreo tipo Gibbs. Para ello se siguen, a grandes rasgos, los algoritmos descritos por Kim y Nelson (1999), para el caso en el que la varianza del componente común es fija, y por Kim, Shepard y Chib (1998), para introducir volatilidad estocástica en dicho componente común. En particular, se siguen, básicamente, tres pasos:

- 1 Se estima el componente común inobservado (f_t, \dots, f_T) condicionando en la volatilidad estocástica del factor ($\sigma_{f_t}, \dots, \sigma_{f_T}$) y en todos los parámetros del modelo (β, σ, ϕ). Este procedimiento se basa en el algoritmo de simulación suavizada propuesto por Carter y Kohn (1994) o Durbin y Koopman (2002).
- 2 En el segundo paso, se estima la volatilidad estocástica del factor ($\sigma_{f_t}, \dots, \sigma_{f_T}$) condicionando en el componente común inobservado (f_t, \dots, f_T) y en todos los parámetros del modelo (β, σ, ϕ). Para ello se sigue la metodología descrita en Kim, Shepard y Chib (1998).
- 3 Por último, condicionando en el componente común inobservado (f_t, \dots, f_T) y en la volatilidad estocástica del factor ($\sigma_{f_t}, \dots, \sigma_{f_T}$), las ecuaciones (1)-(4) son independientes entre sí, lo que permite su tratamiento individual, y la estimación bayesiana de cada uno de los parámetros del modelo se puede realizar de forma estándar (véase Kim and Nelson, 1999).

La identificación del modelo se obtiene asumiendo que tanto el factor de carga asociado a una de las variables (en el caso del modelo descrito se toma como referencia el PIB) como ω_{f_t} en la ecuación (6) son iguales a la unidad.

Anejo 3. Período COVID como no observado

En este anejo se describe una aproximación empírica que consiste en descartar, para la estimación del modelo, el período en el que se producen las mayores oscilaciones en las dinámicas de las variables —en adelante, «modelo *Missing*»—, como se ha sugerido recientemente en la literatura. En este sentido, Maroz, Stock y Watson (2021) consideran un amplio conjunto de indicadores para la economía de Estados Unidos, y afirman que el «período COVID» es aquel que abarca desde marzo hasta junio de 2020. Asimismo, sugieren que existe evidencia a favor de que, a finales de 2020, los indicadores económicos retornaron a sus patrones históricos. Consecuentemente, proponen que una alternativa para la estimación empírica es descartar, para la estimación del modelo, el período en el que se producen las mayores oscilaciones en las dinámicas de las variables. Un ejemplo de la aplicación de esta metodología en Estados Unidos es llevado a cabo por Schorfheide y Song (2021), que encuentran una mejora significativa en sus proyecciones en tiempo real.

En el caso de España, se define como «período COVID» aquel que incluye desde marzo hasta julio de 2020, ya que la variabilidad de los indicadores económicos fue más notable y, a su vez, coincidió con el período en el que se implementaron las restricciones de movilidad más severas. La principal ventaja de esta estrategia empírica es que permite estimar el modelo para los períodos posteriores a la pandemia teniendo en cuenta tanto la información previa a febrero de 2020 como la posterior a julio de 2020¹⁹, sin que las fuertes variaciones observadas en ese período afecten a las relaciones de largo plazo de las variables.

Cuadro A3.1

ERRORES DE PREDICCIÓN EN LOS PERÍODOS PRE- Y POS-COVID

		Pre-COVID (I TR 2015 - IV TR 2019)		Pos-COVID (I TR 2021 - II TR 2023)		Pos-COVID (I TR 2021 - II TR 2023) excluyendo II TR 2021	
		PIB (estimación <i>flash</i>)	PIB (2.ª estimación)	PIB (estimación <i>flash</i>)	PIB (2.ª estimación)	PIB (estimación <i>flash</i>)	PIB (2.ª estimación)
Error medio absoluto	Modelo <i>Anterior</i> *	0,10	0,12	0,67	0,67	0,61	0,67
	Modelo <i>Missing</i>			0,59	0,52	0,44	0,54
	Modelo SV	0,08	0,10	0,25	0,46	0,25	0,35
Raíz del error cuadrático medio	Modelo <i>Anterior</i> *	0,11	0,14	0,85	0,84	0,82	0,86
	Modelo <i>Missing</i>			0,79	0,67	0,50	0,70
	Modelo SV	0,11	0,13	0,30	0,60	0,30	0,41

FUENTES: Elaboración propia.

NOTA: Los errores cuadráticos medios y el error medio absoluto que se muestran están relacionados con las previsiones realizadas en el tercer mes del trimestre objetivo. El modelo *Anterior* se refiere al modelo descrito en el epígrafe 2, pero incluyendo el ISE sin consumidores, el PMI compuesto y el IPI de construcción con una relación contemporánea (y no adelantada) respecto al factor común. El modelo SV se refiere al modelo que incluye volatilidad estocástica en el factor. Para el cómputo del error cuadrático medio y de la raíz del error cuadrático medio de la séptima y octava columna se excluye el error cometido en el segundo trimestre de 2021.

¹⁹ En el caso del PIB también se considera como no observada la información referida al tercer trimestre de 2020.

Desde un punto de vista metodológico, se reemplazan los valores observados para cada uno de los indicadores durante el «período COVID» por observaciones faltantes. Esta alternativa es viable debido a que, como ya se ha mencionado anteriormente, la estimación del modelo se realiza mediante el filtro de Kalman.

Ahora bien, en términos de capacidad predictiva, la evidencia sugiere que la opción *Missing* no supera a la alternativa de incorporar volatilidad estocástica en el factor descrita en el texto principal. Ello se refleja en el cuadro A3.1²⁰, que recoge el error medio absoluto y la raíz del error cuadrático tanto del modelo *Missing* como del modelo *SV*.

²⁰ Nótese que, para el período previo al COVID-19, el modelo *Anterior* y el modelo *Missing* son el mismo modelo, ya que la inclusión de observaciones faltantes en el período de la pandemia solo puede afectar a las previsiones posteriores a dicho período. Por ello, no se presentan los resultados en el período pre-COVID (2015-2019).

PUBLICACIONES DEL BANCO DE ESPAÑA

DOCUMENTOS OCASIONALES

- 2201 MANUEL A. PÉREZ ÁLVAREZ: Nueva asignación de derechos especiales de giro. (Existe una versión en inglés con el mismo número).
- 2202 PILUCA ALVARGONZÁLEZ, MARINA GÓMEZ, CARMEN MARTÍNEZ-CARRASCAL, MYROSLAV PIDKUYKO y ERNESTO VILLANUEVA: Analysis of labor flows and consumption in Spain during COVID-19.
- 2203 MATÍAS LAMAS y SARA ROMANIEGA: Elaboración de un índice de precios para el mercado inmobiliario comercial de España. (Existe una versión en inglés con el mismo número).
- 2204 ÁNGEL IVÁN MORENO BERNAL y TERESA CAMINERO GARCÍA: Analysis of ESG disclosures in Pillar 3 reports. A text mining approach.
- 2205 OLYMPIA BOVER, LAURA CRESPO y SANDRA GARCÍA-URIBE: El endeudamiento de los hogares en la Encuesta Financiera de las Familias y en la Central de Información de Riesgos: un análisis comparativo. (Existe una versión en inglés con el mismo número).
- 2206 EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO y ROBERTO RAMOS: Dinámicas de población durante el COVID-19. (Existe una versión en inglés con el mismo número).
- 2207 JULIO GÁLVEZ: Measuring the equity risk premium with dividend discount models.
- 2208 PILAR CUADRADO, MARIO IZQUIERDO, JOSÉ MANUEL MONTERO, ENRIQUE MORAL-BENITO y JAVIER QUINTANA: El crecimiento potencial de la economía española tras la pandemia. (Existe una versión en inglés con el mismo número).
- 2209 PANA ALVES, SERGIO MAYORDOMO y MANUEL RUIZ-GARCÍA: La financiación empresarial en los mercados de renta fija: la contribución de la política monetaria a mitigar la barrera del tamaño. (Existe una versión en inglés con el mismo número).
- 2210 PABLO BURRIEL, IVÁN KATARYNIUK y JAVIER J. PÉREZ: Computing the EU's SURE interest savings using an extended debt sustainability assessment tool.
- 2211 LAURA ÁLVAREZ, ALBERTO FUERTES, LUIS MOLINA y EMILIO MUÑOZ DE LA PEÑA: La captación de fondos en los mercados internacionales de capitales en 2021. (Existe una versión en inglés con el mismo número).
- 2212 CARLOS SANZ: El peso del sector público en la economía: resumen de la literatura y aplicación al caso español.
- 2213 LEONOR DORMIDO, ISABEL GARRIDO, PILAR L'HOTELLERIE-FALLOIS y JAVIER SANTILLÁN: El cambio climático y la sostenibilidad del crecimiento: iniciativas internacionales y políticas europeas. (Existe una versión en inglés con el mismo número).
- 2214 CARMEN SÁNCHEZ y JARA QUINTANERO: Las empresas *fintech*: panorama, retos e iniciativas.
- 2215 MARÍA ALONSO, EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DIANA POSADA, PATROCINIO TELLO-CASAS y CARLOS TRUCHARTE: La accesibilidad presencial a los servicios bancarios en España: comparación internacional y entre servicios. (Existe una versión en inglés con el mismo número).
- 2216 BEATRIZ GONZÁLEZ, ENRIQUE MORAL-BENITO e ISABEL SOLER: Schumpeter Meets Goldilocks: the Scarring Effects of Firm Destruction.
- 2217 MARIO ALLOZA, JÚLIA BRUNET, VÍCTOR FORTE-CAMPOS, ENRIQUE MORAL-BENITO y JAVIER J. PÉREZ: El gasto público en España desde una perspectiva europea. (Existe una versión en inglés con el mismo número).
- 2218 PABLO AGUILAR, BEATRIZ GONZÁLEZ y SAMUEL HURTADO: Carbon tax sectoral (CATS) model: a sectoral model for energy transition stress test scenarios.
- 2219 ALEJANDRO MUÑOZ-JULVE y ROBERTO RAMOS: Estimación del impacto de variaciones en el período de cálculo de la base reguladora sobre la cuantía de las nuevas pensiones de jubilación. (Existe una versión en inglés con el mismo número).
- 2220 LUIS ÁNGEL MAZA: Una estimación de la huella de carbono en la cartera de préstamos a empresas de las entidades de crédito en España. (Existe una versión en inglés con el mismo número).
- 2221 SUSANA MORENO SÁNCHEZ: The EU-UK relationship: regulatory divergence and the level playing field.
- 2222 ANDRÉS ALONSO-ROBISCO y JOSÉ MANUEL CARBÓ: Inteligencia artificial y finanzas: una alianza estratégica.
- 2223 LUIS FERNÁNDEZ LAFUERZA, MATÍAS LAMAS, JAVIER MENCÍA, IRENE PABLOS y RAQUEL VEGAS: Análisis de la capacidad de uso de los colchones de capital durante la crisis generada por el COVID-19. (Existe una versión en inglés con el mismo número).
- 2224 SONSOLES GALLEGOS, ISABEL GARRIDO e IGNACIO HERNANDO: Las líneas del FMI para aseguramiento y prevención de crisis y su uso en Latinoamérica (Existe una versión en inglés con el mismo número).
- 2301 LAURA HOSPIDO, CARLOS SANZ y ERNESTO VILLANUEVA: Air pollution: a review of its economic effects and policies to mitigate them.
- 2302 IRENE MONASTEROLO, MARÍA J. NIETO y EDO SCHETS: The good, the bad and the hot house world: conceptual underpinnings of the NGFS scenarios and suggestions for improvement.

- 2303 IADRÍAN LÓPEZ GONZÁLEZ: Inteligencia artificial aplicada al control de calidad en la producción de billetes.
- 2304 BELÉN AROCA MOYA: Conceptos, fundamentos y herramientas de neurociencia, y su aplicación al billete.
- 2305 MARÍA ALONSO, EDUARDO GUTIÉRREZ, ENRIQUE MORAL-BENITO, DIANA POSADA y PATROCINIO TELLO-CASAS: Un repaso de las diversas iniciativas desplegadas a nivel nacional e internacional para hacer frente a los riesgos de exclusión financiera.
- 2306 JOSÉ LUIS ROMERO UGARTE, ABEL SÁNCHEZ MARTÍN y CARLOS MARTÍN RODRÍGUEZ: Alternativas a la evolución de la operativa bancaria mayorista en el Eurosistema. (Existe una versión en inglés con el mismo número).
- 2307 HENRIQUE S. BASSO, OURANIA DIMAKOU and MYROSLAV PIDKUYKO: How inflation varies across Spanish households.
- 2308 LAURA CRESPO, NAJIBA EL AMRANI, CARLOS GENTO y ERNESTO VILLANUEVA: Heterogeneidad en el uso de los medios de pago y la banca *online*: un análisis a partir de la Encuesta Financiera de las Familias (2002-2020).
- 2309 HENRIQUE S. BASSO, OURANIA DIMAKOU y MYROSLAV PIDKUYKO: How consumption carbon emission intensity varies across Spanish households.
- 2310 IVÁN AUCIELLO-ESTÉVEZ, JOSEP PIJOAN-MAS, PAU ROLDAN-BLANCO y FEDERICO TAGLIATI: Dual labor markets in Spain: a firm-side perspective.
- 2311 CARLOS PÉREZ MONTES, JORGE E. GALÁN, MARÍA BRU, JULIO GÁLVEZ, ALBERTO GARCÍA, CARLOS GONZÁLEZ, SAMUEL HURTADO, NADIA LAVÍN, EDUARDO PÉREZ ASENJO e IRENE ROIBÁS: Marco de análisis sistémico del impacto de los riesgos económicos y financieros. (Existe una versión en inglés con el mismo número).
- 2312 SERGIO MAYORDOMO e IRENE ROIBÁS: La traslación de los tipos de interés de mercado a los tipos de interés bancarios. (Existe una versión en inglés con el mismo número).
- 2313 CARLOS PÉREZ MONTES, ALEJANDRO FERRER, LAURA ÁLVAREZ ROMÁN, HENRIQUE BASSO, BEATRIZ GONZÁLEZ LÓPEZ, GABRIEL JIMÉNEZ, PEDRO JAVIER MARTÍNEZ-VALERO, SERGIO MAYORDOMO, ÁLVARO MENÉNDEZ PUJADAS, LOLA MORALES, MYROSLAV PIDKUYKO y ÁNGEL VALENTÍN: Marco de análisis individual y sectorial del impacto de los riesgos económicos y financieros. (Existe una versión en inglés con el mismo número).
- 2314 PANA ALVES, CARMEN BROTO, MARÍA GIL y MATÍAS LAMAS: Indicadores de riesgos y vulnerabilidades en el mercado de la vivienda en España.
- 2215 ANDRÉS AZQUETA-GAVALDÓN, MARINA DIAKONOVA, CORINNA GHIRELLI y JAVIER J. PÉREZ: Sources of economic policy uncertainty in the euro area: a ready-to-use database.
- 2316 FERNANDO GARCÍA MARTÍNEZ y MATÍAS PACCE: El sector eléctrico español ante el alza del precio del gas y las medidas públicas en respuesta a dicha alza. (Existe una versión en inglés con el mismo número).
- 2317 ROBERTO BLANCO y SERGIO MAYORDOMO: Evidencia sobre el alcance de los programas de garantías públicas y de ayudas directas a las empresas españolas implementados durante la crisis del COVID-19. (Existe una versión en inglés con el mismo número).
- 2318 ISABEL GARRIDO y IRUNE SOLERA: Has the 2021 general SDR allocation been useful? For what and for whom?
- 2319 ROBERTO BLANCO, ELENA FERNÁNDEZ, MIGUEL GARCÍA-POSADA y SERGIO MAYORDOMO: An estimation of the default probabilities of Spanish non-financial corporations and their application to evaluate public policies.
- 2320 BANCO DE ESPAÑA: La accesibilidad presencial a los servicios bancarios en España: Informe de seguimiento 2023. (Existe una versión en inglés con el mismo número).
- 2321 EDUARDO AGUILAR GARCÍA, MARIO ALLOZA FRUTOS, TAMARA DE LA MATA, ENRIQUE MORAL-BENITO, IÑIGO PORTILLO PAMPIN y DAVID SARASA FLORES: Una primera caracterización de las empresas receptoras de fondos NGEU en España.
- 2401 ALEJANDRO MORALES, MANUEL ORTEGA, JOAQUÍN RIVERO y SUSANA SALA: ¿Cómo identificar a todas las sociedades del mundo? La experiencia del código LEI (Legal Entity Identifier).
- 2402 XAVIER SERRA y SONSOLES GALLEGÓ: Un primer balance del *Resilience and Sustainability Trust* del FMI como canal de utilización de los derechos especiales de giro.
- 2403 PABLO HERNÁNDEZ DE COS: El papel de la política macroprudencial en la estabilización de las fluctuaciones macrofinancieras. Conferencia de Estabilidad Financiera/Banco de Portugal, Lisboa (Portugal), 2 de octubre de 2023.
- 2404 MORTEZA GHOMI, SAMUEL HURTADO y JOSÉ MANUEL MONTERO: Análisis de la dinámica reciente de la inflación en España. Un enfoque basado en el modelo de Blanchard y Bernanke (2023).
- 2405 PILUCA ALVARGONZÁLEZ, MARINA ASENSIO, CRISTINA BARCELÓ, OLYMPIA BOVER, LUCÍA COBREROS, LAURA CRESPO, NAJIBA EL AMRANI, SANDRA GARCÍA-URIBE, CARLOS GENTO, MARINA GÓMEZ, PALOMA URCELAY, ERNESTO VILLANUEVA and ELENA VOZMEDIANO: The Spanish Survey of Household Finances (EFF): description and methods of the 2020 wave.
- 2406 ANA GÓMEZ LOSCOS, MIGUEL ÁNGEL GONZÁLEZ SIMÓN y MATÍAS JOSÉ PACCE: Modelo para la previsión del PIB de la economía española a corto plazo en tiempo real (Spain-STING): nueva especificación y reevaluación de su capacidad predictiva.