

Big Data y consumo: el índice BBVA de comercio al por menor⁸

Los avances recientes en el almacenamiento, gestión y procesamiento de datos han dado lugar a un incremento cuantitativo y cualitativo de la información disponible para el análisis económico, tanto desde la perspectiva individual como desde la óptica macroeconómica. **El desarrollo de la tecnología Big Data ha posibilitado el análisis en tiempo real de la información** generada por ciudadanos, gobiernos y empresas en todas aquellas interacciones que dejan una huella digital. Nuevas fuentes de datos, como las originadas por el uso de redes sociales y los motores de búsqueda en Internet, se han usado, entre otros fines, para la predicción de la evolución de variables económicas como el empleo, el consumo o los flujos turísticos (Chamberlin, 2010, McLaren, 2011, Choi y Varian, 2012, entre otros). Más recientemente, Cavallo (2016) y Cavallo y Rigobon (2016) han demostrado que los precios de bienes y servicios ofrecidos en Internet pueden ser utilizados para calcular índices de alta frecuencia que eviten sesgos de medición y permitan analizar precios relativos en una muestra amplia de países.

En reconocimiento a la utilidad de esta nueva disciplina, instituciones como Eurostat han creado grupos de trabajo en torno al paradigma *Big Data*. Por otro lado, se hace patente que los datos generados en el marco de la actividad de empresas privadas con una amplia base de clientes complementan las fuentes recabadas por la

administración pública. En esta dirección, Sobolevsky *et al.* (2015) han demostrado la posibilidad de obtener señales socioeconómicas a nivel regional en España a través del gasto realizado mediante tarjetas de crédito y débito.

Para extraer el valor analítico que encierran los datos generados por la interacción de los clientes en los sistemas de prestación de servicios electrónicos de BBVA, se ha establecido una colaboración entre dos áreas con grandes sinergias en sus capacidades: BBVA Research y BBVA Data & Analytics. Los primeros resultados de esta línea de investigación se exponen en el **presente recuadro, en el que se estudia el comportamiento agregado de los consumidores mediante la información de las transacciones realizadas con tarjeta de crédito o débito por los clientes de BBVA en terminales punto de venta (TPV) españoles.** A partir de esta fuente de información, se construye un indicador del gasto en comercios minoristas (ICM-BBVA) y se presentan sus desagregaciones temporal, geográfica y sectorial, que superan en granularidad y frecuencia a las de las fuentes estadísticas oficiales.

El ICM-BBVA replica la estructura del índice de comercio al por menor publicado mensualmente por el Instituto Nacional de Estadística (ICM-INE). Esta variable es fundamental a la hora de estudiar no sólo la evolución del consumo privado⁹, sino también la del ciclo económico¹⁰. Como ilustra el Gráfico R.1.1, el ICM-INE es un

8: Este recuadro sintetiza la investigación realizada por BBVA Data & Analytics y BBVA Research (de próxima publicación).

9: El consumo privado representó el 57% del PIB español en 2015.

10: En numerosos estudios sobre la evolución del ciclo económico de EE. UU., el ICM surge como uno de los cuatro indicadores fundamentales junto con el empleo, los ingresos y el índice de producción (véase, por ejemplo, Stock y Watson, 1989).

Recuadro1

indicador adelantado del gasto de los hogares en España, lo que potencia las ventajas de contar con un índice que, como el ICM-BBVA, reproduce su comportamiento y puede llegar a estar disponible en tiempo real y con mayor desagregación.

Los resultados de este trabajo indican que el Big Data tiene un elevado potencial para enriquecer el estudio sobre la toma de decisiones de los agentes económicos. Se demuestra la factibilidad de replicar fielmente la evolución de macromagnitudes relevantes como el ICM-INE, no sólo a nivel nacional, sino también por comunidades autónomas (CC. AA.) y modos de distribución¹¹. Además, la mayor granularidad que ofrece esta fuente de datos respecto a la proporcionada por las estadísticas oficiales permite explorar con detalle la evolución del gasto de los hogares. Por ejemplo, se detectan patrones regulares de consumo (semanales, mensuales y anuales) y se cuantifica la influencia de los festivos en el gasto.

El índice BBVA de comercio al por menor (ICM-BBVA)

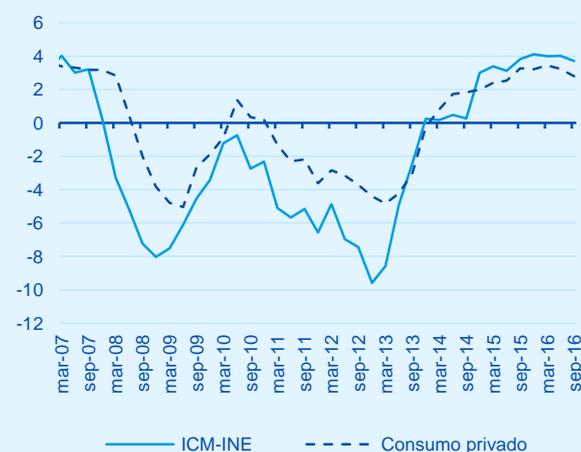
Las limitaciones inherentes a la recolección y procesamiento de la información dentro del análisis tradicional (por ejemplo, el coste y la frecuencia) provocan que la mayoría de los indicadores socioeconómicos publicados por los institutos de estadística se basen en encuestas por muestreo en lugar de censos poblacionales. Los estadísticos incluidos en el ICM-INE no son la excepción. Tienen por objeto informar de la evolución reciente de la

11: Además de las estaciones de servicio, el ICM-INE distingue cuatro modos de distribución según la dimensión del establecimiento: grandes superficies, grandes cadenas, pequeñas cadenas y empresas unilocalizadas.

actividad del comercio minorista en España, y para ello, se recoge información relativa al empleo y la facturación de 12.000 establecimientos del sector¹². Los indicadores se publican mensualmente con un retardo de un mes respecto al periodo de referencia, se desglosan a nivel de CC. AA. y, para el agregado nacional, por modos de distribución¹³.

Gráfico R.1.1

España: ventas al por menor y consumo privado (datos CVEC, % a/a)



Fuente: BBVA Research a partir de INE

Los avances ocurridos durante los últimos años en el acopio, almacenamiento y gestión de grandes conjuntos de datos, junto con la creciente bancarización y digitalización de las economías, potencian la capacidad de análisis económico en tiempo real. En particular, el registro de las transacciones realizadas mediante tarjetas de crédito y

12: Todas aquellas empresas cuya actividad principal se englobe dentro de la división 47 de la Clasificación Nacional de Actividades Económicas (CNAE-2009), comercio al por menor, excepto de vehículos de motor y motocicletas. En consecuencia, el ICM-INE no incluye el gasto en servicios de restauración y hostelería.

13: Los ICM-INE se calculan según un índice de Laspeyres encadenado y tienen como base el año 2010. Para más información, véase la nota metodológica disponible en: http://www.ine.es/daco/daco43/notaccm10_cnae2009.pdf

Recuadro1

débito¹⁴ puede complementar y ampliar las señales que se extraen de la facturación del comercio al por menor estimada por el INE. Para contrastar esta hipótesis, se utiliza información anonimizada del gasto con tarjetas bancarias realizado por clientes de BBVA en establecimientos dedicados a la venta al por menor en España¹⁵ y se construye el indicador de consumo ICM-BBVA.

El ICM-BBVA replica la estructura sectorial del ICM-INE y amplía sus dimensiones temporal y geográfica mediante un minucioso procedimiento de selección, extracción, categorización y depuración de la información (véase el Gráfico R.1.2). En primer lugar, se seleccionan las transacciones en establecimientos dedicados a la venta al por menor de bienes de consumo. A continuación, se categorizan las observaciones en los distintos modos de distribución contemplados en el ICM-INE. En tercer lugar, se depuran los datos ante la posible existencia de duplicidades, errores u omisiones y se excluyen aquellas transacciones que, por su cuantía y localización, puedan considerarse atípicas. Finalmente, se agrega la información y se vuelve a realizar el procedimiento de forma iterativa hasta que se verifica la idoneidad estadística de los indicadores obtenidos respecto a las fuentes oficiales.

14: La información de la transacción incluye, entre otras características, su importe, geolocalización y sector de distribución, así como atributos descriptivos del comprador que en ningún caso posibilitan la identificación del usuario de la tarjeta.

15: La muestra total alcanza los cuatro millones de clientes, que suponen un total de 200 millones de transacciones al año.

Gráfico R.1.2

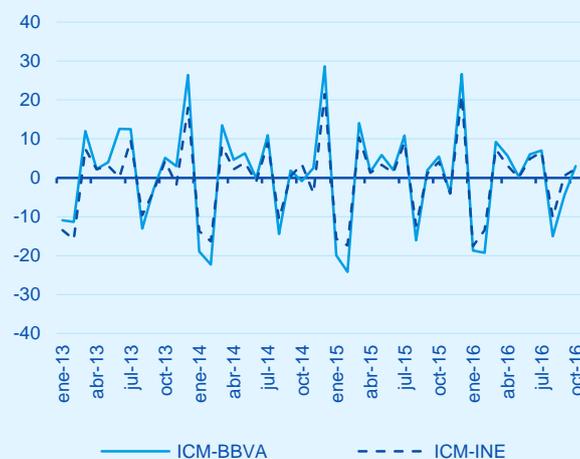
Proceso de extracción y tratamiento de los datos de BBVA



Fuente: BBVA Data & Analytics

Gráfico R.1.3

Ventas al por menor vs ICM-BBVA (datos nominales brutos, % m/m)



Fuente: BBVA a partir de INE

Los resultados obtenidos validan el procedimiento llevado a cabo para la construcción de los indicadores. Como se observa en el Gráfico R.1.3, las tasas mensuales de crecimiento del ICM-INE¹⁶ y del ICM-BBVA para el conjunto de España presentan una dinámica muy similar. Éste resultado se mantiene cuando se

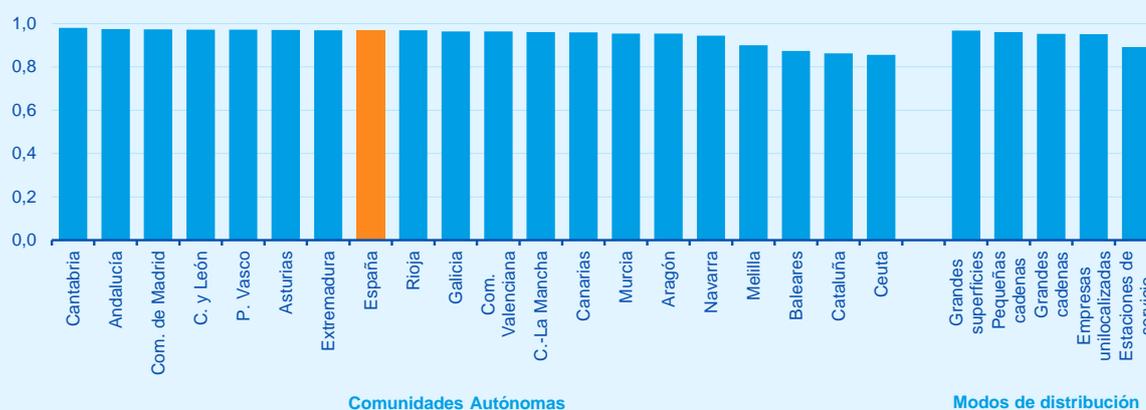
16: Índice nominal y sin corregir de variaciones estacionales y efecto calendario.

desagregan los índices por CC. AA. o por modos de distribución. De hecho, cuando se compararan las tasas de crecimiento mensual de los índices publicados por el INE y las construidos a partir de la información

de BBVA se aprecia que la correlación contemporánea supera el 0,85 en todos los casos (véase el Gráfico R.1.4).

Gráfico R.1.4

Correlación entre las variaciones mensuales de las ventas al por menor y del ICM-BBVA (datos nominales brutos, ene-13/oct-16)



Fuente: BBVA a partir de INE

Ventajas del índice BBVA de comercio al por menor: frecuencia y granularidad

El nivel de detalle de la información utilizada para la construcción del ICM-BBVA (localización, fecha y actividad del establecimiento) permite la confección de índices que, a partir de la metodología descrita en el apartado anterior, ofrecen un mayor nivel de desagregación que el proporcionado por las estadísticas oficiales. Para dar a conocer las potencialidades del ICM-BBVA, se presentan a continuación algunos resultados sobre la evolución del comercio minorista a nivel provincial, sectorial y diario.

La geolocalización de los TPV permite obtener los ICM-BBVA a nivel provincial.

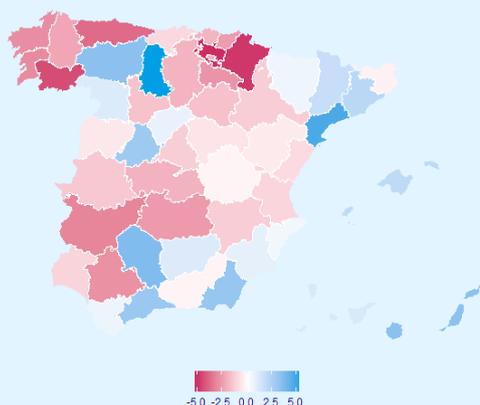
Si bien no es factible contrastar la dinámica de estas series con las estadísticas del INE

(la información del ICM-INE se publica por CC. AA.), la consistencia en los niveles de agregación superior, junto con la correspondencia entre el ICM-INE y el ICM-BBVA para las CC. AA. uniprovinciales, confieren una verosimilitud elevada al resto de índices provinciales. En el Gráfico R.1.5 se observa, por ejemplo, cómo se ha comportado el comercio minorista en cada una de las provincias respecto a la media española. Así, al comparar el crecimiento interanual del gasto acumulado hasta septiembre de 2016, se aprecia que Palencia, Tarragona, Córdoba, Las Palmas y León han sido las provincias con mejor desempeño del comercio al por menor en España. En contraste, la facturación en las provincias del norte y del centro de la península ha crecido significativamente

menos que la media durante el año en curso¹⁷.

Gráfico R.1.5

**ICM-BBVA provincial
(diferencial de crecimiento nominal acumulado en 2016 respecto a la media nacional, en pp)**

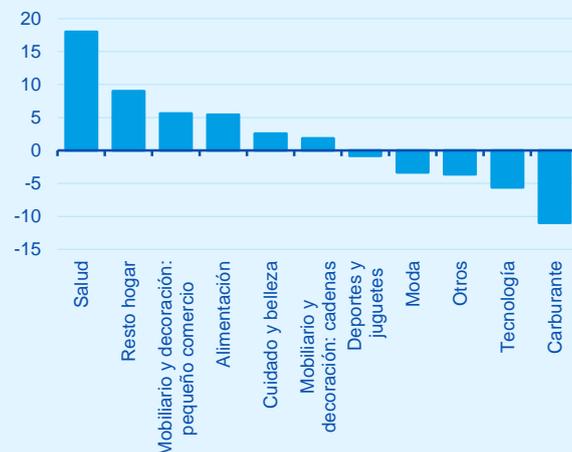


Fuente: BBVA a partir de INE

Del análisis del **ICM-BBVA por sectores** se desprende que el consumo de bienes de primera necesidad (alimentación y salud), de cuidado personal (droguería, perfumería y peluquería y estética) y de equipamiento del hogar (mobiliario y decoración, bricolaje, jardinería y floristería, entre otros) ha crecido por encima de la media durante el último año. Por el contrario, el gasto nominal en moda, tecnología y carburante ha exhibido un dinamismo comparativamente reducido, como muestra el Gráfico R.1.6.

Gráfico R.1.6

**ICM-BBVA sectorial
(diferencial de crecimiento nominal acumulado en 2016 respecto a la media nacional, en pp)**



Fuente: BBVA

Por último, el cómputo del **ICM-BBVA diario** ofrece la posibilidad de profundizar en el estudio de los patrones de consumo de los agentes, imposibles de desentrañar con un índice mensual. En particular, la mayor frecuencia no sólo permite analizar la estacionalidad a lo largo del año, sino también cuál es el comportamiento dentro de una semana o de un mes, e incluso cómo afectan los días festivos a los hábitos de consumo¹⁸.

Los resultados principales de la modelización se sintetizan en los Gráficos R.1.7 y R.1.8. El primero describe el perfil de consumo dentro de la semana y evidencia que **los sábados concentran más ventas minoristas respecto a la media semanal, mientras que lo opuesto sucede los**

17: Los resultados son cualitativamente similares a los analizados en el Observatorio Regional del 16 de noviembre, disponible en: <https://www.bbva-research.com/publicaciones/espana-dinamismo-y-elevada-dispersion-regional-en-2016-que-se-reduciran-en-2017/>

18: Para llevar a cabo este análisis, se ha empleado una descomposición estructural del ICM-BBVA diario nacional siguiendo la metodología propuesta por Harvey, Koopman y Riani (1997). El modelo incluye una tendencia estocástica, variables dicotómicas que capturan el efecto "día de la semana" e introduce *splines* para modelizar la estacionalidad dentro del mes y el año. El efecto de los días festivos se captura con variables binarias que no sólo afectan al festivo en sí mismo, sino también a algunos días anteriores y posteriores. Para más detalles, véase BBVA (2017).

Recuadro1

domingos¹⁹. El segundo de los gráficos ilustra el patrón de consumo medio dentro del mes. **A pesar de que el gasto mensual es relativamente estable, se aprecia un repunte a comienzos y finales de mes que podría corresponderse con el momento de cobro de la nómina.** En cuanto al análisis del patrón del consumo dentro del año, los resultados se encuentran en línea con lo esperado dada la estacionalidad de las series mensuales publicadas por el INE (mayor consumo en julio y diciembre). Finalmente, **se constata que los festivos afectan positivamente al gasto en los días previos y posteriores a la celebración del mismo y, como era de esperar, negativamente en el día en cuestión.**

Gráfico R.1.7

ICM-BBVA diario: distribución semanal del consumo (% del gasto nominal en cada día de la semana, 1ª semana de junio)

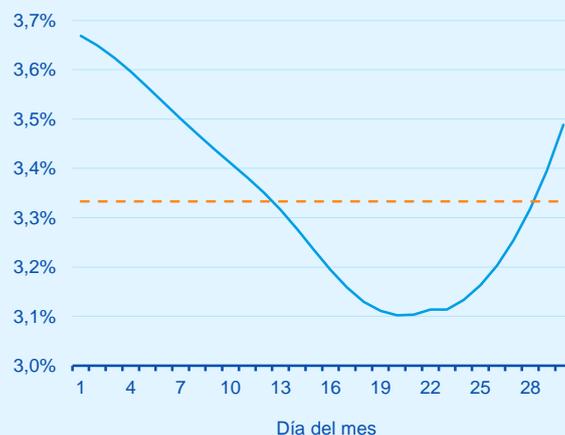


Fuente: BBVA a partir de INE

19: Este comportamiento se debe, al menos en parte, a las restricciones de apertura comercial aún vigentes en algunas Comunidades Autónomas.

Gráfico R.1.8

ICM-BBVA diario: distribución mensual del consumo (% del gasto nominal en cada día del mes)



Fuente: BBVA a partir de INE

Conclusiones

El uso de grandes conjuntos de datos en el análisis económico se ha consolidado durante los últimos años. El presente trabajo contribuye a esta línea de investigación al demostrar la utilidad de la tecnología *Big Data* para replicar la evolución de macromagnitudes relevantes como los indicadores de comercio al por menor. Además, dada su elevada granularidad y frecuencia, permite enriquecer el análisis de las decisiones de los agentes económicos.

Este recuadro demuestra que la evolución de las estadísticas referidas a la facturación del comercio al por menor publicadas por el INE puede estimarse a partir de la información de transacciones con tarjeta de crédito y débito (ICM-BBVA). Dada la riqueza y el volumen de la información, es posible desagregar el índice con un mayor detalle espacial, sectorial y temporal que el publicado en las estadísticas oficiales. En consecuencia, se proponen indicadores que miden la evolución del comercio al por

Recuadro1

menor a nivel provincial, para sectores específicos o con una mayor frecuencia que la oficial. Respecto a este último punto, la estimación de un índice diario permite analizar los patrones regulares que describen el comportamiento de los consumidores dentro de la semana, el mes y el año, así como la influencia de los festivos en el mismo.

Referencias

BBVA (2017). "Computing retail trade in Spain using Big Data", mimeo.

Cavallo, A. (2016). "Are Online and Offline Prices Similar? Evidence from Large Multi-Channel Retailers", *American Economic Review* (Forthcoming).

Cavallo, A. & Rigobon, R. (2016). "The Billion Prices Project: Using Online Prices for Measurement and Research", *Journal of Economic Perspectives*, Spring 2016, Vol 30(2): 151-78.

Chamberlin, G. (2010). "Googling the present", *Economic and Labour Market Review* 4: 59-95.

Choi, H., & Varian, H. (2012). "Predicting present with google trends", *Economic Record* 88: 2-9.

Harvey, A., Koopman, S. J., & Riani, M. (1997). "The modeling and seasonal adjustment of weekly observations", *Journal of Business & Economic Statistics*, 15(3), 354-368.

McLaren, N. (2011). "Using internet search data as economic indicators", *Bank of England Quarterly Bulletin* 51: 134-140.

Sobolevsky, S., Massaro, E., Bojic, I., Arias, J. M., & Ratti, C. (2015). "Predicting regional economic indices using big data of individual bank card transactions", *arXiv preprint arXiv:1506.00036*.

Stock, J. H., & Watson, M. W. (1989). "New indexes of coincident and leading economic indicators", en *NBER Macroeconomics Annual 1989, Volume 4* (pp. 351-409). MIT press.

AVISO LEGAL

El presente documento, elaborado por el Departamento de BBVA Research, tiene carácter divulgativo y contiene datos, opiniones o estimaciones referidas a la fecha del mismo, de elaboración propia o procedentes o basadas en fuentes que consideramos fiables, sin que hayan sido objeto de verificación independiente por BBVA. BBVA, por tanto, no ofrece garantía, expresa o implícita, en cuanto a su precisión, integridad o corrección.

Las estimaciones que este documento puede contener han sido realizadas conforme a metodologías generalmente aceptadas y deben tomarse como tales, es decir, como previsiones o proyecciones. La evolución histórica de las variables económicas (positiva o negativa) no garantiza una evolución equivalente en el futuro.

El contenido de este documento está sujeto a cambios sin previo aviso en función, por ejemplo, del contexto económico o las fluctuaciones del mercado. BBVA no asume compromiso alguno de actualizar dicho contenido o comunicar esos cambios.

BBVA no asume responsabilidad alguna por cualquier pérdida, directa o indirecta, que pudiera resultar del uso de este documento o de su contenido.

Ni el presente documento, ni su contenido, constituyen una oferta, invitación o solicitud para adquirir, desinvertir u obtener interés alguno en activos o instrumentos financieros, ni pueden servir de base para ningún contrato, compromiso o decisión de ningún tipo.

Especialmente en lo que se refiere a la inversión en activos financieros que pudieran estar relacionados con las variables económicas que este documento puede desarrollar, los lectores deben ser conscientes de que en ningún caso deben tomar este documento como base para tomar sus decisiones de inversión y que las personas o entidades que potencialmente les puedan ofrecer productos de inversión serán las obligadas legalmente a proporcionarles toda la información que necesiten para esta toma de decisión.

El contenido del presente documento está protegido por la legislación de propiedad intelectual. Queda expresamente prohibida su reproducción, transformación, distribución, comunicación pública, puesta a disposición, extracción, reutilización, reenvío o la utilización de cualquier naturaleza, por cualquier medio o procedimiento, salvo en los casos en que esté legalmente permitido o sea autorizado expresamente por BBVA.

Este informe ha sido elaborado por la unidad de España y Portugal

Economista Jefe de España y Portugal

Miguel Cardoso
miguel.cardoso@bbva.com
+34 91 374 39 61

Ignacio Archondo
ignacio.archondo@bbva.com
+34 91 757 52 78

Joseba Barandiaran
joseba.barandia@bbva.com
+34 94 487 67 39

Álvaro Flores
alvaro.flores.alonso@bbva.com
+34 91 757 52 78

Con la colaboración de:

Sistemas Financieros
Macarena Ruesta
esperanza.ruesta@bbva.com

Juan Ramón García
juanramon.gl@bbva.com
+34 91 374 33 39

Félix Lores
felix.lores@bbva.com
+34 91 374 01 82

Antonio Marín
antonio.marin.campos@bbva.com
+34 648 600 596

Turquía, China y Geopolítica
Tomasa Rodrigo
tomasa.rodrigo@bbva.com

Myriam Montañez
miriam.montanez@bbva.com
+34 638 80 85 04

Matías José Pacce
matias.pacce@bbva.com
+34 647 392 673

Virginia Pou
virginia.pou@bbva.com
+34 91 537 77 23

Juan Ruiz
juan.ruiz2@bbva.com
+34 646 825 405

Pep Ruiz
ruiz.aguirre@bbva.com
+34 91 537 55 67

Camilo Andrés Ulloa
camiloandres.ulloa@bbva.com
+34 91 537 84 73

BBVA Research

Economista Jefe Grupo BBVA

Jorge Sicilia Serrano

Análisis Macroeconómico

Rafael Doménech
r.domenech@bbva.com

Escenarios Económicos Globales

Miguel Jiménez
mjimenezg@bbva.com

Mercados Financieros Globales

Sonsoles Castillo
s.castillo@bbva.com

Modelización y Análisis de Largo Plazo Global

Julián Cubero
juan.cubero@bbva.com

Innovación y Procesos

Oscar de las Peñas
oscar.delaspenas@bbva.com

Sistemas Financieros y Regulación

Santiago Fernández de Lis
sfernandezdelis@bbva.com

Coordinación entre Países

Olga Cerqueira
olga.gouveia@bbva.com

Regulación Digital

Álvaro Martín
alvaro.martin@bbva.com

Regulación

María Abascal
maria.abascal@bbva.com

Sistemas Financieros

Ana Rubio
arubiog@bbva.com

Inclusión Financiera

David Tuesta
david.tuesta@bbva.com

España y Portugal

Miguel Cardoso
miguel.cardoso@bbva.com

Estados Unidos

Nathaniel Karp
Nathaniel.Karp@bbva.com

México

Carlos Serrano
carlos.serrano@bbva.com

Turquía, China y Geopolítica

Álvaro Ortiz
alvaro.ortiz@bbva.com

Turquía

Álvaro Ortiz
alvaro.ortiz@bbva.com

China

Le Xia
le.xia@bbva.com

América del Sur

Juan Manuel Ruiz
juan.ruiz@bbva.com

Argentina

Gloria Sorensen
gsorensen@bbva.com

Chile

Jorge Selaive
jselaive@bbva.com

Colombia

Juana Téllez
juana.tellez@bbva.com

Perú

Hugo Perea
hperea@bbva.com

Venezuela

Julio Pineda
juliocesar.pineda@bbva.com