

ANÁLISIS MACROECONÓMICO

Perú | Búsquedas en internet y proyecciones de ventas de departamentos en Lima

Vanessa Belapatiño y Yalina Crispin

En este informe se presenta un índice de interés de compra de departamentos para la ciudad de Lima (IICDG), el que se ha elaborado empleando información proporcionada por *Google Trend*. Se encuentra evidencia de que el IICDG anticipa (con cinco trimestres) y que mejora la capacidad para predecir las transacciones en el sector inmobiliario. Finalmente, empleando el IICDG en un modelo de proyección para las ventas de departamentos se anticipa que estas se mantendrán estables durante el 2017.

Búsquedas en internet, *Google Trend* y el análisis del comportamiento de los agentes económicos

Las búsquedas en internet contienen información relevante sobre las preferencias e intenciones de millones de compradores y empresas, y su digitalización está permitiendo su uso sistemático. *Google*, uno de los principales buscadores de información en internet a nivel mundial, ha tenido un papel clave en este proceso al canalizar y consolidar los distintos datos disponibles en diversas páginas web para facilitar el acceso de todos los usuarios.

Una herramienta gratuita de *Google* es *Google Trend*, la cual proporciona información sobre la cantidad de búsquedas para una frase o palabra. Esta información se brinda mediante un índice de frecuencia semanal y es posible acotarla a una región geográfica particular. El índice se construye sobre la base de dos procesos. El primero consiste en la normalización: la cantidad de búsquedas de la palabra o frase se divide entre el total búsquedas hechas en *Google*, controlando por área geográfica e intervalo de tiempo que representa. El objetivo de este proceso es aislar a los datos de los efectos del crecimiento del número de usuarios de internet y del aumento de la popularidad de *Google* como buscador¹, por lo que resulta un indicador de interés. El segundo proceso corresponde al escalamiento: se asume con valor de 100 el número de búsquedas más alto en todo el periodo de análisis, por lo que el índice está acotado a valores entre 0 y 100.

Dos características importantes sobre la información que brinda *Google Trend*: es en tiempo real y de alta frecuencia, lo que permite obtener datos actualizados de la consulta, a diferencia de otras fuentes que brindan información mensual o trimestral y con cierto periodo de retraso.

Estudios recientes demuestran que los datos de búsqueda en internet son útiles para realizar proyecciones en diversas disciplinas

En los últimos años, diversos estudios han mostrado la efectividad del uso de los datos de *Google* para mejorar las proyecciones en los campos de la epidemiología y de la economía. Sobre la primera disciplina, entre los trabajos más conocidos están los de Polgreen et al (2008) y Ginsberg et al (2009), los que mostraron que los datos de búsqueda en internet ayudan a predecir la incidencia de enfermedades como la influenza. Estos trabajos tuvieron amplia difusión y estimularon hallazgos adicionales en epidemiología.

¹ Carrie – Swallow y Labbé (2010).

En el campo de la economía, Choi y Varian (2009 y 2011) describen cómo utilizar las búsquedas de *Google* para predecir variables económicas como las ventas minoristas, ventas de automóviles, viviendas y viajes². Los autores encontraron que los modelos que incorporan como regresor al índice de *Google Trend* mejoran su capacidad predictiva.

Otros autores han examinado el uso de información de internet para el análisis del mercado laboral. Por ejemplo, Askitas y Zimmerman (2010) emplean información de internet para predecir el desempleo en Alemania. También se pueden mencionar dentro de esta línea el trabajo de Baker y Fradkim (2011) sobre la dinámica de empleo en Estados Unidos y el de Chang y Río (2013) que estima un modelo de predicción del nivel de empleo para Perú.

Por otro lado, Pena y Huang (2009) construyeron un índice de confianza del consumidor para Estados Unidos sobre la base de los cambios en los patrones de búsqueda en *Google*. Encontraron que el índice construido está altamente correlacionado con el índice de confianza del consumidor de la Universidad de Michigan (ICS) y con el índice de confianza del consumidor del Conference Board (CCI). Los autores, además, encuentran que el índice de *Google* anticipa los cambios de los dos índices de referencia. Asimismo, encuentran que el índice de *Google* tiende a superar tanto al ICS como al CCI para predecir el crecimiento del gasto de consumo personal y de las ventas minoristas.

Otros trabajos que usan información de búsqueda en internet están relacionados al mercado de vivienda. Wu y Brynjolfsson (2009) construyen un índice de búsquedas de vivienda en Estados Unidos que les permite predecir las ventas del mercado inmobiliario y los precios. En Askitas y Zimmermann (2011) se encuentra evidencia de la relación entre la morosidad hipotecaria y la búsqueda en internet de las solicitudes de modificación de préstamos (un prestatario le escribe a su banco cuando tiene inconvenientes y externaliza su deterioro económico).

Como se puede apreciar, el uso de la información de búsquedas en internet tiene diversas aplicaciones y un amplio potencial de desarrollo para diversas disciplinas como la economía.

Elaboración de un índice de Interés de Compra de Departamentos para Lima con información del *Google Trends* (IICDG)

Para la construcción del IICDG para Lima se ha procedido de la siguiente manera:

1. Para la búsqueda en internet, se define una lista de frases relacionadas con el interés de compra de un departamento. En concreto, se seleccionaron cinco frases que mostraron la mayor representatividad³: (i) “departamentos venta Lima”, (ii) “departamentos en venta en Lima”, (iii) “venta de departamentos en Lima”, (iv) “inmobiliaria en Lima” y (v) “inmobiliarias en Lima”.
2. Las series elegidas se transforman –mediante promedios simples– de frecuencia mensual a trimestral⁴. Esta transformación se realiza para igualar la frecuencia del IICDG con la de ventas de departamentos de Lima (que sólo está disponible trimestralmente)⁵, para luego analizar si existe una relación entre estas dos variables.
3. Se ponderaron las cinco series por la inversa de sus desviaciones estándar para reducir la importancia de sus volatilidades y la incertidumbre en el IICDG (más detalle ver Anexo 1).

² Hal Varian es Economista Principal de Google Inc.

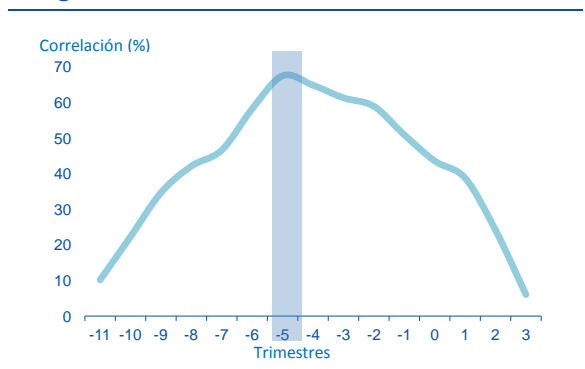
³ Debido a que Google Trend muestra la búsqueda de palabras o frases exactas, algunas de las frases seleccionadas son similares

⁴ Se consideró como dato del tercer trimestre del 2016 al promedio de los meses de julio y agosto.

⁵ Se consideró la serie de ventas de departamento en Lima elaborada por TINSA (información disponible a partir del primer trimestre del 2008).

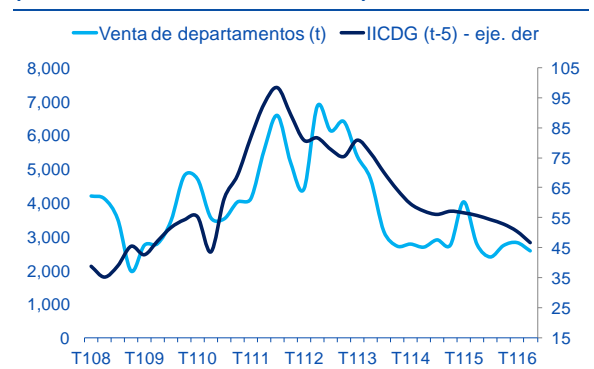
Como se puede apreciar en el Gráfico 1, el análisis de correlación dinámica revela que existe una relación positiva entre el IICDG y la venta de departamentos⁶, alcanzando su mayor asociación cuando el IICDG se adelanta en cinco trimestres (ver Gráfico 1). Es así que, al graficar las venta de departamentos y el IICDG con un rezago de cinco trimestres⁷, se muestra un comportamiento similar de ambas series (ver Gráfico 2). Además, al realizar pruebas precedencia estadística⁸, se encontró que el IICDG anticipa (causa) el comportamiento de las ventas de departamentos en Lima⁹ (ver Anexo 2).

Gráfico 1
Coefficiente de correlación entre las Ventas de departamentos en Lima y el IICDG con diferentes rezagos



Fuente: TINSA, Google
Elaboración: BBVA Research

Gráfico 2
Venta de departamentos en Lima e IICDG (número de unidades e índice*)



*En el primer trimestre del 2010 se registró el máximo nivel de búsquedas, por ello el índice toma el valor de 100.
Fuente: TINSA, Google
Elaboración: BBVA Research

Con lo anterior, encontramos evidencia de que el interés de compra de departamentos medido por el IICDG se anticipa a la realización de la venta. Este resultado es coherente en la práctica debido a que la compra de un departamento requiere un tiempo para evaluar: la ubicación, el precio, características al gusto del comprador, a qué inmobiliaria acudir, el tiempo de ahorro para la cuota inicial y con qué entidad endeudarse, etc.

Adicional al tiempo de búsqueda de un departamento, debe sumarse el que le toma a la entidad bancaria aprobar un crédito hipotecario. La evaluación no toma mucho tiempo si el comprador es un trabajador dependiente de ingresos medios, pero no siempre es así. Existe otro proceso de evaluación más largo y se utiliza cuando el trabajador es independiente. En este caso, debido a que el comprador no puede garantizar un ingreso permanente, el sistema bancario le brinda la posibilidad de acceder a un crédito si mantiene un depósito constante de por lo menos 6 meses.

En ese contexto, se analiza la relación entre el índice de interés de compra de un departamento (IICDG) y el crecimiento del crédito hipotecario. Se encuentra que el IICDG también se anticipa a la dinámica del crédito hipotecario. Al rezagar los créditos hipotecarios en 4 trimestres, se encuentra la correlación más alta (ver Gráfico 3 y 4). Además, al realizar pruebas de precedencia

⁶ Las dos series se encuentran desestacionalizadas.

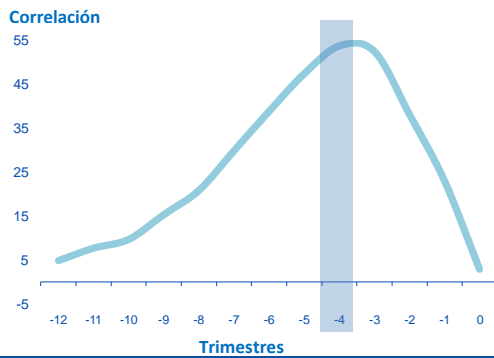
⁷ Por ejemplo, el dato de venta de departamento del segundo trimestre del 2016 se relacionado con el valor del IICDG del primer trimestre del 2015; es decir, el valor del IICDG se anticipa al comportamiento de las ventas de departamentos.

⁸ Pruebas de causalidad a lo Granger. Ver detalles en Anexo 2.

⁹No obstante, es importante resaltar que la afirmación “IICDG causa a las ventas de departamento en Lima” no implica que las ventas de departamento en Lima sea el efecto o el resultado del IICDG, pues intervienen, además otros factores.

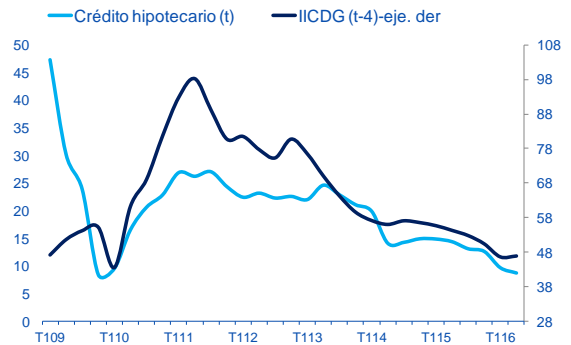
estadística se encontró que el IICDG anticipa (causa) el crecimiento de los créditos hipotecarios (ver Anexo 2).

Gráfico 3
Coeficiente de correlación entre el crecimiento del crédito hipotecario del sistema bancario y el IICDG



Fuente: Google, ASBANC
 Elaboración: BBVA Research

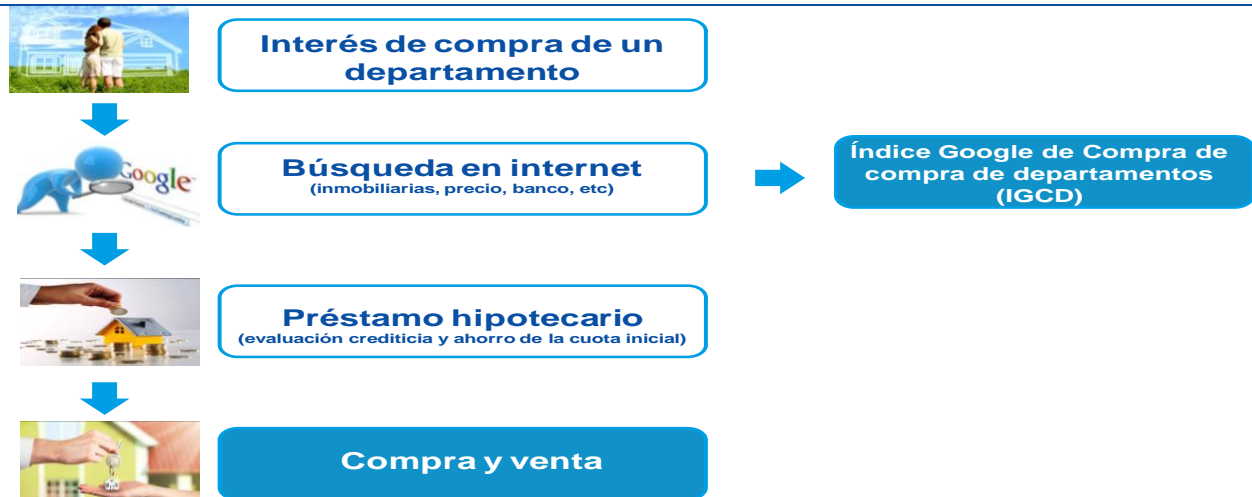
Gráfico 4
Crédito hipotecario del sistema bancario y el IGCD (var.% interanual e Índice*)



*En el primer trimestre del 2010 se registró el máximo nivel de búsqueda, por ello el índice toma el valor de 100.
 Fuente: ASBANC y Google
 Elaboración: BBVA Research

En síntesis, los resultados encontrados muestran evidencia de la relación que existe entre las búsquedas en internet, que refleja el interés de las familias en comprar un departamento, y la venta de este (la realización de la transacción).

Gráfico 5
Relación entre interés de las familias por comprar un departamento (medido por el IGCD) y la compra y venta



Elaboración: BBVA Research

El Índice de interés de compra de departamentos en Lima (IICDG) mejora la capacidad para predecir las ventas de departamentos

Para evaluar si el uso del IICDG mejora la capacidad predictiva de las ventas de departamentos, se estimaron dos modelos econométricos (Ver Anexo 3): uno con especificación ARMA (Modelo de Referencia) y otro que incluye información el IICDG (Modelo Aumentado). Se encontró que la capacidad predictiva del segundo modelo es mejor que el de referencia. Como se puede apreciar en la Tabla 1, el modelo que incluye el IICDG presenta un menor error cuadrático medio (ECM) y un menor error medio absoluto (EMA) que el Modelo de Referencia. Además, el R^2 aumenta ligeramente en el Modelo Aumentado. Asimismo, cuando se realiza la estimación fuera de la muestra, en el modelo Aumentado tiene menores ECM y EMA en 46% y 52%, respectivamente, con respecto al Modelo de Referencia.

Tabla 1

Indicadores de capacidad predictiva de los dos modelos de proyección

Indicadores	Modelo de referencia	Modelo aumentado
R-squared	81	83
Dentro de la muestra		
ECM ^{1/}	575	539
EMA	464	385
Fuera de de la muestra		
ECM ^{1/}	1,485	804
EMA	1,276	612

*1/ Se consideró la raíz del error cuadrático medio (ECM).
Elaboración: BBVA Research

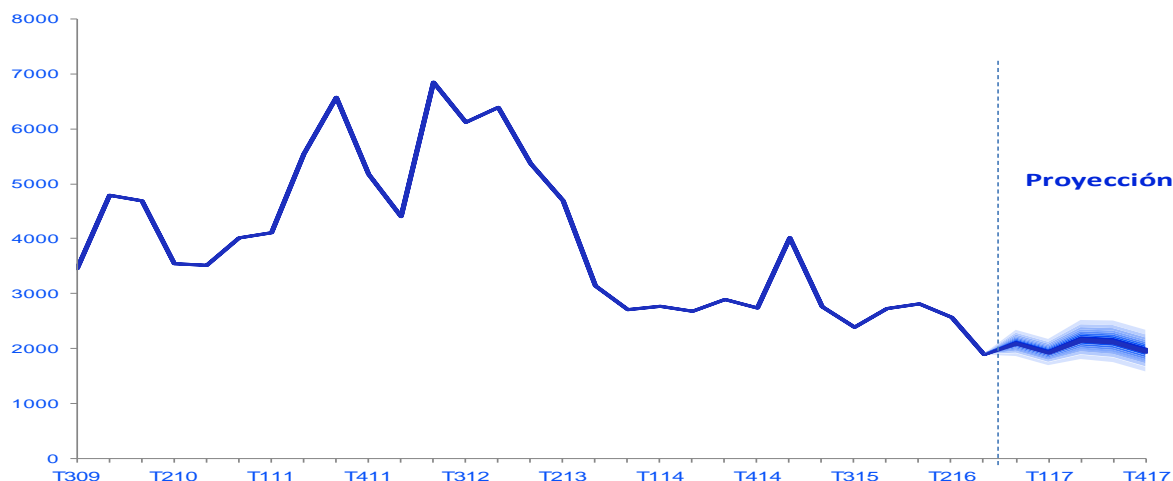
El nivel que ha mostrado el IICDG en lo que va del año es consistente con un nivel de ventas estable durante el 2017

El interés por la compra de departamentos, medido por el IICDG, detuvo su desaceleración en el tercer trimestre del 2015 y ha permanecido estable hasta agosto del presente año. Considerando que el lapso de tiempo que transcurre entre el interés y la compra de es de cinco trimestres (como se evidencia en las secciones anteriores), la evolución en lo que va del año del IICDG sugiere que el nivel de venta de departamentos se mantendrá relativamente estable en lo que resta del 2016 y 2017.

El Modelo de proyección que incluye el IICDG (descrito en la sección anterior) también apunta a que la venta de departamentos en Lima detendrá su desaceleración y mantendrá niveles similares a los actuales en los próximos meses (ver Grafico 9).

Gráfico 9

Venta de departamento en Lima Metropolitana: resultados del modelo aumentado * (número de unidades)



*Proyección fuera de la muestra.

Fuente: Tinsa

Elaboración: BBVA Research

Cabe agregar que algunos eventos recientes, que aún no están siendo capturados por el IICDG, podrían dar un mayor impulso al mercado inmobiliario en los siguientes meses. Por ejemplo, los cambios normativos para los fondos de pensiones (retiros para la compra de nueva vivienda¹⁰).

Conclusiones

La información de búsquedas en internet permite tener una previsión de corto plazo del mercado inmobiliario. Para esto, se construyó un indicador de interés de compra de departamentos (IICDG), a través de *Google Trend*, el cual anticipa, con cinco trimestres, el comportamiento de sus ventas futuras.

El Índice de interés de compra de departamentos en Lima (IICDG) mejora la capacidad para predecir las ventas de departamentos Se planteó una estimación de dos modelos econométricos: uno con especificación ARMA (Modelo de Referencia) y otro que incluye información del IICDG (Modelo Aumentado). Se encontró que la capacidad predictiva del modelo aumentado es mejor que el de referencia.

El nivel que ha mostrado el IICDG en lo que va del año es consistente con un nivel de ventas estable durante el 2017. El interés por la compra de departamentos detuvo su desaceleración en el tercer trimestre del 2015 y ha permanecido estable hasta agosto del presente año. Considerando que el lapso de tiempo que transcurre entre el interés y la compra de es de cinco trimestres, el nivel de venta de departamentos se mantendrá estable en lo que resta del 2016 y 2017. Cabe agregar que algunos eventos podrían impulsar el mercado inmobiliario, como es el caso de la liberalización de los fondos de las AFPs para la adquisición de vivienda.

¹⁰ En Junio de 2016 se promulgó la Ley de AFPs que impulsa la compra de una vivienda. La Ley N° 30478 permite el retiro hasta del 25% del fondo de los afiliados a una AFP, con la condición de que dicho dinero se destine a la cuota inicial para la compra de un vivienda o para adelantar pago de un crédito hipotecario, ambos para el caso de una primera vivienda. Se estima que alrededor de 1 millón de personas se verán beneficiadas de la ley.

Anexo 1: Construcción del Índice Google de Compra de Departamento

Para la construcción del Índice Google de Compra de Departamento (IGCD), se indexaron las cinco series obtenidas de las búsquedas en Google Trend de las siguientes frases: “departamentos venta lima”, “departamentos en venta en lima”, “venta de departamentos en lima”, “inmobiliaria en lima” y “inmobiliarias en lima”

El primer paso para la indexación fue calcular la desviación estándar de cada una de las series que representan a cada una de las cinco frases. Ver las ecuaciones 1 y 2, donde σ_n representa la desviación de n , siendo n una de las 5 series y t uno de los trimestres.

$$\sigma_n = \sqrt{\frac{\sum_{t=1}^T (x_{tn} - \bar{x}_n)^2}{T - 1}} \quad \forall \quad t = 1 a T, \quad n = 1 a N \quad (1)$$

$$\text{donde} \quad \bar{x}_n = \frac{\sum_{t=1}^T x_{tn}}{T} \quad (2)$$

Luego, se calcula a cada ponderador como la división de la inversa de la desviación estándar de cada serie entre la suma de las inversas de las desviaciones estándar (ecuación 3).

$$\rho_n = \frac{\frac{1}{\sigma_i}}{\sum_{n=1}^N \frac{1}{\sigma_i}} \quad (3)$$

A partir de ello, se obtiene de la suma producto de las cinco series iniciales y sus respectivos ponderadores (ecuación 4).

$$\text{Índice_ponderado}_t = \sum_{n=1}^N \rho_n x_{tn} \quad (4)$$

Finalmente, para la construcción del IGCD se cambia la escala a la serie calculada en la ecuación 4. Para ello, se valora como 100 al trimestre que muestra el mayor interés por la compra de un departamento. Así, el primer trimestre del 2010 toma el valor de 100 y los otros se recalculan conforma a la ecuación 4.

Anexo 2: Prueba de causalidad a lo Granger ¹¹

Causalidad a lo Granger: IICDG y Venta de departamentos en Lima*

Hipótesis nula	Probabilidad
a. IICDG no causa a las ventas de departamento	0.06
b. Ventas de departamento no causa IICDG	0.83

*Con una probabilidad de 10% se rechaza la hipótesis nula a. Es decir el IICDG es una variable que permite predecir de las ventas de departamentos en Lima (se hizo la prueba con un rezago y cinco rezagos y se rechaza la hipótesis nula a.)

Causalidad a lo Granger: IICDG y el crecimiento del crédito hipotecario *

Hipótesis nula	Probabilidad
a. IICDG no causa a créditos	0.00
b. créditos no causa IICDG	0.11

*Con una probabilidad de 5% se rechaza la hipótesis nula a. Es decir el IICDG es una variable que permite realizar predecir el crecimiento del crédito hipotecario (se hizo la prueba con un rezago y cuatro rezagos y se rechaza la hipótesis nula a.)

¹¹ La prueba de Granger analiza la relación de causalidad entre dos variables, evaluación distinta a calcular solo correlaciones. Esta prueba permite afirmar si una variable "y" es causada por otra "x", es decir, si "x" ayuda a estimar y predecir "y". Para ello se evalúa si los coeficientes de la variable "x" rezagada son estadísticamente significativos al estimar "y". Sin embargo, la afirmación "x" causa a "y" no implica que "y" sea el resultado de "x", ya que intervienen, además, otros factores al margen de "x" que afectan a "y". En ocasiones, se puede encontrar una doble causalidad-Granger: "x" causa a "y" e "y" causa a "x".

Anexo 3: Modelos de proyección

Parámetros estimados, variable endógena: Venta de departamentos en Lima (y_t)*

Variables exógenas	Modelo de referencia	Modelo aumentado
Yt(-1)	0.73	0.47
t- st.	5.7	2.89
MA(12)	-0.91	0.90
t- st.	29.4	29.40
Constante	4,016	-2,533
t- st.	8.5	-2.5
IGCD**		94.64
t- st.		6.6

*Las especificación del modelo ARMA consideró la metodología Box- Jenkins (las variables son estacionarias y los modelos no presentan autocorrelación ni heterocedasticidad). ** La variable se encuentra rezagada en cinco trimestres.
Elaboración: BBVA Research

Prueba de Raíz Unitaria para la serie IGCD y venta de departamentos

Null Hypothesis: IGCD_SA has a unit root
Exogenous: Constant
Lag Length: 7 (Automatic - based on SIC, maxlag=8)

	t-Statistic
Elliott-Rothenberg-Stock DF-GLS test statistic	-4.185157
Test critical values:	
1% level	-2.653401
5% level	-1.953858
10% level	-1.609571

*MacKinnon (1996)

Null Hypothesis: VENTA_SA has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=8)

	t-Statistic
Elliott-Rothenberg-Stock DF-GLS test statistic	-1.947061

Test critical values:	1% level	-2.636901
	5% level	-1.951332
	10% level	-1.610747

*MacKinnon (1996)

Prueba: Breush Godfrey de autocorrelación

Modelo de referencia

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.380924	Prob. F(2,28)	0.6867
Obs*R-squared	0.420578	Prob. Chi-Square(2)	0.8103

Modelo Aumentado

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.293380	Prob. F(2,27)	0.7481
Obs*R-squared	0.241060	Prob. Chi-Square(2)	0.8865

Prueba: White de heterocedasticidad

Modelo Base

Heteroskedasticity Test: White

F-statistic	1.749012	Prob. F(9,23)	0.1344
Obs*R-squared	13.40841	Prob. Chi-Square(9)	0.1450
Scaled explained SS	11.08604	Prob. Chi-Square(9)	0.2699

Modelo Aumentado

Heteroskedasticity Test: White

F-statistic	2.770427	Prob. F(4,28)	0.05
Obs*R-squar	9.357225	Prob. Chi-Square(4)	0.05
Scaled expla	14.81991	Prob. Chi-Square(4)	0.01

Referencias

1. Carrière-Swallow, y. y F. Labbé (2010). Nowcasting with Google Trends in an emerging market, Documento de Trabajo, 2010-588, Banco Central de Chile.
2. Choi Hyunyoung y Hal Varian. Predicting the Present with Google Trends. Google Inc. 2009
3. Choi Hyunyoung y Hal Varian. Predicting the Present with Google Trends. Google Inc. 2011
4. Della Penna, N. y H. Huang (2009). Constructing Consumer Sentiment Index for U.S. Using Google Searches. Universidad de Alberta, Canadá. 2009.
5. Jeremy Ginsberg, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. Detecting influenza epidemics using search engine query data. 2009
6. Jillie Chang y Andrea Del Río. Google Trends: Predicción del nivel de empleo agregado en Perú usando datos en tiempo real, 2005-2011. Working Paper BCRP. 2013
7. Lynn Wu y Erik Brynjolfsson. How Google Searches Foreshadow Housing Prices and Sales. 2009
8. Nikos Askitas and Klaus F. Zimmermann. Detecting Mortgage Delinquencies with Google Trends. 2010.
9. Nikos Askitas and Klaus F. Zimmermann. Google econometrics and unemployment forecasting. 2010.
10. Scott Baker and Andrey Fradkin. What Drives Job Search? Evidence from Google Search Data. Stanford Institute for Economic Policy Research. 2011
11. P. M. Polgreen, y. Chen, D. M. Pennock, and F. D. Nelson. Using internet searches for influenza surveillance. Clinical Infectious Diseases. 2008.

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.