



Dinámica inflacionaria regional y el esquema de metas de inflación en el Perú

DIEGO WINKELRIED Y JOSÉ ENRIQUE GUTIÉRREZ*

La política monetaria del Banco Central de Reserva del Perú (BCRP) se ha basado en un esquema de metas de inflación por más de una década, utilizando la inflación de Lima Metropolitana como medida operacional. Un indicador alternativo es la inflación nacional, cuya calidad estadística y disponibilidad en tiempo real han mejorado sustancialmente en los últimos años. Dados estos indicadores alternativos cabe preguntarse: ¿Cuáles han sido las implicancias para la inflación nacional de haber monitoreado la inflación en Lima? ¿Monitorear la inflación nacional afectaría significativamente el funcionamiento de la política monetaria en el Perú? Para responder a estas preguntas, se estima un modelo de corrección de errores entre las inflaciones regionales y se investiga cómo se propagan los choques inflacionarios a lo largo del país. El modelo incorpora (i) restricciones de agregación, donde cada inflación regional es afectada por un agregado de las inflaciones de las regiones vecinas y (ii) restricciones de largo plazo que permiten identificar una tendencia común en el sistema. Los resultados indican que un choque en la inflación de Lima se transmite rápidamente a las inflaciones en provincias, y explica gran parte de su variabilidad. Con ello, se concluye que al monitorear la inflación en Lima, el BCRP ha contribuido con la estabilidad de la inflación nacional.

Palabras Clave : Inflación regional, PPC relativa, modelo de corrección de errores.

Clasificación JEL : C32, C50, E31, E52, R10.

La política monetaria en el Perú se ha basado en un esquema de metas de inflación por casi dos décadas. A principios de los años 90, el Banco Central de Reserva del Perú (BCRP) se embarcó en un programa de desinflación que redujo la inflación desde niveles hiperinflacionarios (alrededor de 140 por ciento en 1991) a niveles de un dígito (6.5 por ciento en 1997) y, finalmente, a niveles internacionales (3.7 por ciento en 1999). El proceso de desinflación fue gradual mientras el BCRP acumulaba la experiencia necesaria para monitorear la inflación. Con el fin de anclar las expectativas de inflación, en 1994 el BCRP empezó

* Winkelried: Departamento de Modelos Macroeconómicos, Banco Central de Reserva del Perú, Jr. Antonio Miró Quesada 441, Lima 1, Perú. Teléfono: +511 613-2000 (e-mail: diego.winkelried@bcrp.gob.pe). Gutiérrez: Superintendencia de Banca, Seguros y Administradoras Privadas de Fondos de Pensiones (SBS), Los Laureles 214, Lima 27, Perú. Teléfono: +511 630-9000 (e-mail: jgutierrez@sbs.gob.pe).

Agradecemos a Gustavo Yamada, Marco Ortiz, Adrián Armas, Gabriel Rodríguez, participantes de seminarios de investigación en el BCRP y la SBS, participantes del XXX Encuentro de Economistas del BCRP, y a un árbitro anónimo por sus útiles comentarios. Estamos también agradecidos con Peter Paz, Brenda Pizarro y Jimena Montoya por su asistencia en la investigación. Las opiniones en este documento no reflejan necesariamente la visión del BCRP o de la SBS.

a anunciar metas de inflación anuales y en 2002 adoptó completamente un régimen de metas explícitas inflación, donde se anunciaron objetivos de largo plazo (véase Rossini y Vega, 2008).

La medida operacional de inflación utilizada por el BCRP ha sido siempre la variación del Índice de Precios al Consumidor (IPC) de la capital peruana, Lima. Existieron buenas razones para esta elección. En primer lugar, la economía peruana está fuertemente centralizada: Lima concentra casi un tercio de población nacional y solía representar más del 70 por ciento del gasto nacional. Por tanto, para todo efecto práctico, parece razonable pensar que la inflación de Lima puede servir como un buen indicador de la inflación nacional (véase Armas y otros, 2010). En segundo lugar, a pesar de que los datos del IPC nacional están disponibles, la fecha de su publicación solía plantear dificultades para el seguimiento del estado de la economía en tiempo real. Históricamente, mientras que la cifra definitiva del IPC de Lima se publica el primer día hábil del mes posterior, la publicación del IPC nacional ha estado sujeta a retrasos e incluso a revisiones.

Sin embargo, esta situación parece haber cambiado. En la última década, la economía peruana ha crecido a una importante tasa anual promedio de 6.5 por ciento, y muchas provincias han crecido a un ritmo aún más rápido. Aunque el país se mantiene centralizado y Lima representa 2/3 del gasto nacional, la importancia económica de algunas provincias se ha incrementado significativamente y así se ha ido reduciendo la brecha de ingresos con respecto a la capital. Por otro lado, después de mejoras en el diseño de encuestas y técnicas de muestreo, desde inicios de 2012 el Instituto Nacional de Estadística e Informática (INEI) se ha comprometido a publicar la inflación nacional de una forma tan confiable y oportuna como la de Lima. Estos desarrollos han llevado a la academia local y a la prensa económica a indagar acerca de la idoneidad de la inflación nacional, en vez de la de Lima, como objetivo de la política monetaria.¹ Surgen dos interesantes preguntas de política:

- ¿Cuáles han sido las implicancias para la inflación nacional de que el BCRP haya implementado metas inflacionarias con relación a Lima?
- ¿Migrar hacia un agregado nacional afectaría significativamente el funcionamiento de la política monetaria en el Perú?

Con el fin de responder a estas preguntas, es necesario cuantificar cuán relevante es el comportamiento de la inflación en Lima en la determinación de la inflación en el resto del país. Para ello se toma a la paridad del poder de compra (PPC) como marco conceptual, propicio para el estudio de las diferencias en los precios entre regiones de un país o en una unión monetaria (Rogers, 2007). En su versión absoluta, la PPC predice que, en ausencia de costos de transacción y otras fricciones, los precios en dos regiones deberían converger a un mismo *nivel* (véase Parsley y Wei, 1996; Cecchetti y otros, 2002, para estudios exhaustivos). Para el caso peruano, Monge y Winkelried (2009) encuentran que la PPC absoluta se cumple en cerca de 3/4 de todos los posibles pares de ciudades. La convergencia de los precios estaría limitada por los costos de transporte que implica una geografía adversa y la falta de una infraestructura de comunicaciones adecuada en algunas de las regiones del Perú (especialmente en la Sierra).

No obstante, esta noción de PPC no es del todo útil para el análisis de política monetaria, donde el interés se centra en el comportamiento de los diferenciales de *inflación* más que en los diferenciales de

¹ Experiencias similares se pueden encontrar explorando las páginas web de las oficinas de estadísticas y los bancos centrales de otros países de América Latina con esquemas de metas de inflación. En Colombia, el IPC cubrió 13 ciudades desde 1989 hasta 2008, y 24 ciudades a partir de 2009; en México, 35 ciudades hasta febrero de 1995, y 46 ciudades desde marzo de 1995; en Chile, hasta enero de 2009 el IPC fue el de Santiago de Chile, luego en febrero de 2009 los centros urbanos de Puente Alto y San Bernardo se incorporaron, y desde febrero de 2010 el IPC cubre todo el país. No obstante, la representatividad geográfica del IPC es un tema que no está sujeto a mucha discusión en el análisis de política monetaria y, por tanto, no hemos encontrado investigación relacionada a las implicancias de estos cambios en la medición de la inflación.

precios. Por ello, este estudio se enfoca en una versión más débil, la PPC relativa, que establece que una vez que los efectos de los choques regionales transitorios se disipan, las tasas de inflación de dos ciudades deberían converger a una tasa de equilibrio, incluso si los choques conllevan a discrepancias permanentes en los niveles de precios (véase, por ejemplo, Buseti y otros, 2006; Beck y otros, 2009).

El objetivo de este trabajo es responder a las preguntas antes planteadas mediante el estudio de las relaciones dinámicas entre las inflaciones regionales en Perú, y de cómo éstas podrían afectar el diseño de la política monetaria. En particular, se investiga si la PPC relativa se cumple entre las regiones peruanas y, si ese fuera el caso, si el ajuste hacia el equilibrio es lo suficientemente rápido. Para ello, se estima y simula un modelo de corrección de errores para las tasas de inflación de 9 regiones económicas del Perú, un sistema dinámico que muestra cómo se propagan los choques regionales en todo el país. Para lograr una especificación parsimoniosa pero dinámicamente relevante, se imponen restricciones que limitan los efectos de retroalimentación de otras regiones al considerar agregados del “resto del país” en vez de cada región por separado. Estos agregados se calculan utilizando tanto un enfoque geográfico, donde la influencia de las regiones vecinas es mayor que la de regiones más distantes, y un enfoque económico, donde la importancia de una región es proporcional a su participación en el gasto nacional. Una vez calculados estos agregados, se pueden analizar las relaciones dinámicas entre las tasas de inflaciones regionales y, en particular, se puede determinar si los datos responden a tendencias comunes de inflación.

El incremento de la globalización en las últimas décadas ha puesto en manifiesto la relevancia de los vínculos internacionales en la economía mundial, y la importancia de capturar estos vínculos en modelos empíricos (véase Winkelried y Saldarriaga, 2012). Como resultado, existe una amplia literatura que explota esas interrelaciones, donde el denominado VAR global (GVAR) de Pesaran y otros (2004) es el enfoque más popular. Bajo la óptica de un GVAR, agregados del “resto del mundo” son calculados usando ponderaciones de comercio internacional. El enfoque de modelación en este trabajo está relacionado con el GVAR, pero existen al menos dos diferencias metodológicas importantes. En primer lugar, el modelo acá plantado es más pequeño, ya que incluye una variable por unidad (inflación regional). Aunque ello impide etiquetar a los choques más adecuadamente (por ejemplo, choques de oferta versus choques de demanda), permite contrastar formalmente la hipótesis de agregación que se da por sentada en la literatura GVAR. En segundo lugar, la estrategia de identificación es distinta ya que acá se infieren los choques estructurales regionales relevantes (por ejemplo, un choque a la inflación de Lima) a partir de las propiedades de largo plazo del sistema.

Se presenta evidencia empírica concluyente sobre la importancia de los choques a la inflación de Lima para la fijación de precios en el resto del país. En particular, la mayor parte de la variabilidad de las tasas de inflación regionales es determinada por una tendencia común que se puede identificar a partir de la evolución de la inflación de Lima. Además, se encuentra que la convergencia hacia esta tendencia es rápida, cómodamente dentro del rezago de política monetaria (es decir, el tiempo necesario para que una acción de política monetaria afecte a la inflación, el cual se estima entre uno y dos años). Dada la presencia de mecanismos de corrección de desequilibrios tan activos, se concluye que al enfocarse en la inflación de Lima, el BCRP ha promovido indirectamente la estabilidad de la inflación en todo el país.

El resto del documento está organizado de la siguiente manera. La sección 1 motiva la discusión y presenta algunos hechos estilizados. La sección 2 discute cuestiones metodológicas y desarrolla un modelo de corrección de errores que permite retroalimentaciones parsimoniosamente. Adicionalmente, se proponen contrastes de hipótesis formales sobre las restricciones de agregación y sobre la PPC relativa. La sección 3 presenta los resultados de las estimaciones y analiza los efectos de choques originados tanto en Lima como en el resto del país. La sección 4 presenta las conclusiones y plantea algunos temas para futuras investigaciones. Varios detalles técnicos se relegan a los anexos.

1 UN VISTAZO A LOS DATOS

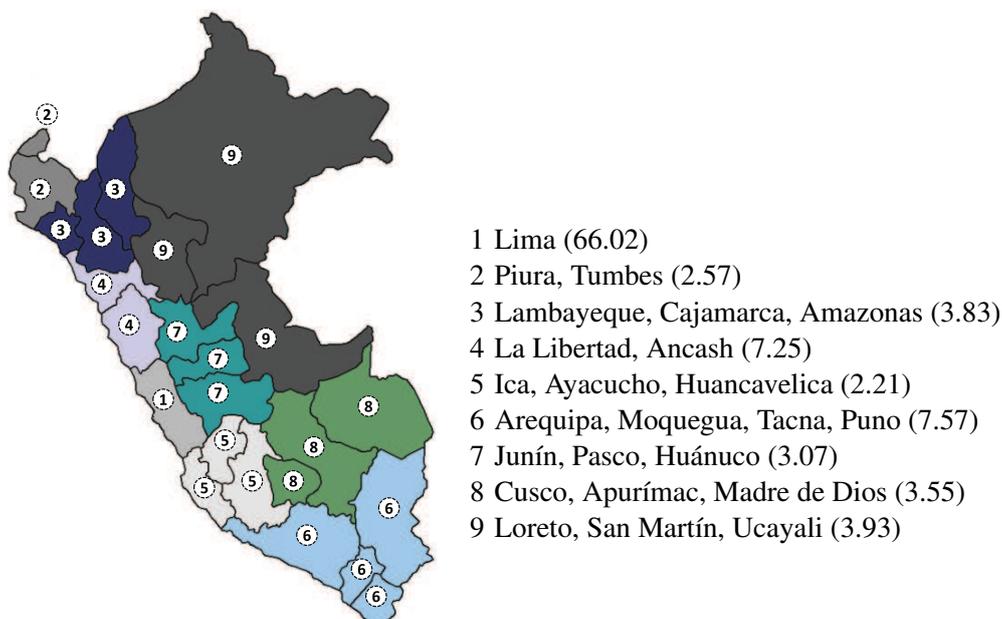
El INEI publica mensualmente datos del IPC de las 25 ciudades más grandes de los 24 Departamentos del Perú. Para trabajar con un número manejable de unidades, se utiliza la clasificación del país en $n = 9$ regiones económicas propuesta por Gonzales de Olarte (2003). Esta clasificación se basa tanto en consideraciones históricas y geográficas (las regiones están formadas por departamentos contiguos) como por motivos económicos, por ejemplo la articulación e integración de mercados.

Las regiones económicas se presentan en el Gráfico 1, junto con las ponderaciones utilizadas por el INEI en el cálculo de la inflación nacional, en base a encuestas de gasto. Como se mencionó, Lima es por mucho la región más importante del país, con una participación de 66.02 por ciento. Es seguida por los centros económicos situados en las regiones 4 (La Libertad y Ancash) y 6 (Arequipa, Moquegua, Tacna y Puno), cada una con una participación ligeramente superior al 7 por ciento. La proporción de cada una de las 6 regiones restantes varía entre 2.21 (región 5) y 3.93 (región 9) por ciento. Vale la pena mencionar que estas participaciones correlacionan fuertemente con la distribución de la población en todo el país.

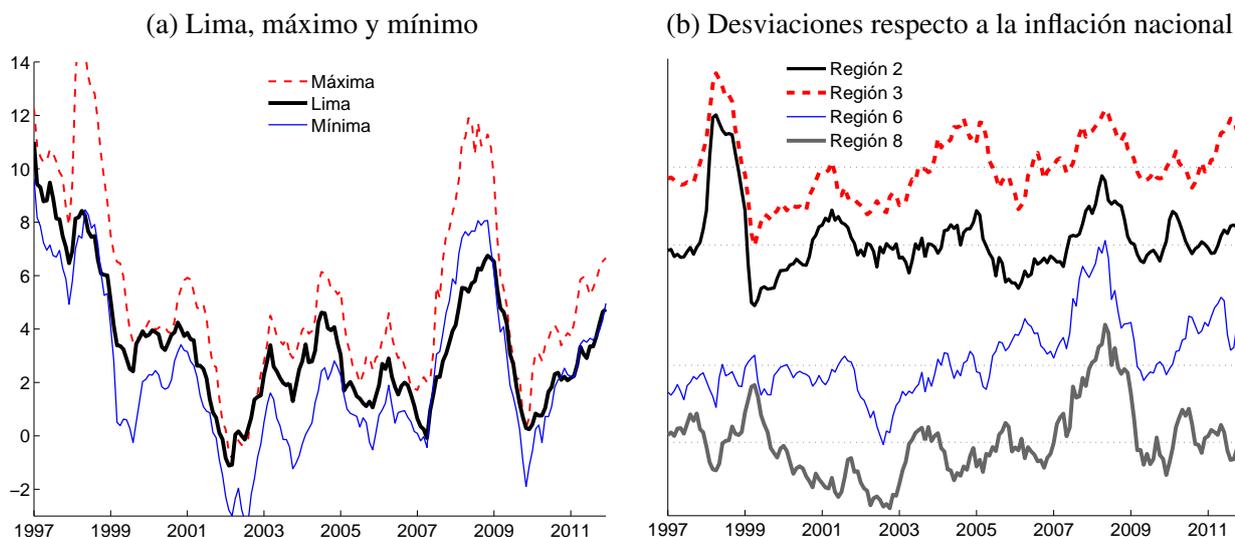
Los datos corresponden al periodo de 1996 a 2011, aproximadamente 190 observaciones después de ajustar por condiciones iniciales. Durante todo el periodo de la muestra, el BCRP ha anunciado metas de inflación. En el análisis empírico más adelante, la inflación se define como el cambio porcentual mensual anualizado de cada IPC P_t , $1200 \log(P_t/P_{t-1})$. Sin embargo, para facilitar la visualización, en esta sección se presentan tasas interanuales (más suaves), $100 \log(P_t/P_{t-12})$.

El panel (a) del Gráfico 2 (p. 83) muestra la evolución de la inflación de Lima, junto con las tasas máximas y mínimas encontradas entre las 8 regiones restantes. Dada la importancia de Lima, a esta escala es difícil distinguir la inflación nacional de la inflación de Lima (con excepción de algunos episodios), por lo que no se incluye en el gráfico. Es notorio que las tres series de inflación presentan muchas características en común. En particular, tienden a moverse en conjunto (sus puntos de inflexión son prácticamente los mismos con pequeñas diferencias de aproximadamente uno o dos meses) y sus

GRÁFICO 1. *Regiones económicas en el Perú*



NOTA: Clasificación de los Departamentos del Perú en 9 regiones económicas, según Gonzales de Olarte (2003, p. 41). Los números en paréntesis son las ponderaciones de cada región al componer la inflación nacional, según el INEI.

GRÁFICO 2. Inflación regional (cambios porcentuales interanuales), 1997 a 2011

NOTAS: Para un periodo dado, el panel (a) muestra la inflación de Lima y la tasa máxima y mínima entre las 8 regiones restantes. En el panel (b), constantes fueron añadidas a las series para facilitar la visualización. Para cada serie, el eje vertical va de -2.5 a cerca de 4.5 por ciento, y las líneas horizontales corresponden a cero.

variaciones son de magnitudes comparables. Por tanto, a primera vista las tasas de inflación parecen estar impulsadas por una tendencia común o, en otras palabras, la PPC relativa parece ser una característica de los datos, un punto que se trata de manera formal más adelante.

Existen, por supuesto, episodios de discrepancia que, aunque importantes, parecen ser temporales. Dos de ellos se destacan. El primero se produce en 1998, cuando el país se vio afectado por un Fenómeno de El Niño particularmente intenso. Los costos económicos de El Niño se relacionan con interrupciones en la producción agrícola y la pesca, y con daños directos a la infraestructura debido a inundaciones, típicamente concentrados en las regiones del norte del país más cercanas al Ecuador (2, 3 y 4). Estas regiones experimentaron un pico de inflación que se revirtió cuando las condiciones climáticas volvieron a la normalidad en 1999.

El segundo episodio corresponde al período de 2007 hacia mediados de 2008 y es consecuencia de dos fenómenos económicos relacionados. El primero es el auge de los precios de *commodities* en los mercados internacionales, principalmente provocado por el impulso de demanda en grandes economías emergentes como China. La magnitud y persistencia de los choques afectaron los precios regionales de manera desigual, principalmente porque la participación de los alimentos en el IPC es mayor en provincias, hecho que explica en parte por qué Lima registró la menor tasa de inflación durante este periodo.

Por otro lado, cifras preliminares del INEI sobre la producción regional indican que durante el período comprendido entre 2005 y 2008, Lima creció a una tasa de 7.5 por ciento anual, mientras que regiones como Arequipa, Cusco y La Libertad crecieron a tasas más cercanas al 8 por ciento, e incluso lugares más exitosos como Ica crecieron por encima del 9 por ciento. Esta fuerte expansión en la producción regional (y la demanda) se vio impulsada por las condiciones extraordinariamente favorables que el país gozó como exportador de materias primas. Ello da un indicio de que muchas regiones han reducido la brecha de ingresos con respecto a Lima de manera relativamente rápida, generando inevitablemente presiones inflacionarias a nivel regional. Adicionalmente, evidencia anecdótica sugiere que las presiones inflacionarias podrían haber sido magnificadas por condiciones iniciales, una suerte de efecto acercamiento (*catch-up*). En particular, los precios inicialmente más bajos en las provincias (debido

al efecto Balassa-Samuelson) habrían crecido a niveles cercanos a los de la capital conforme las provincias se enriquecieron, produciendo así una mayor inflación. Sin embargo, las presiones generadas por estos procesos se desvanecieron rápidamente después de que la crisis financiera internacional comenzara a finales de 2008, tras el colapso del banco de inversión Lehman-Brothers.

El Gráfico 2(b) muestra las desviaciones de la inflación con respecto a la tasa nacional para regiones seleccionadas de los dos extremos del país. Se puede observar que las desviaciones de inflación de la región 2 en el norte están más estrechamente relacionadas con las de la región vecina 3 que con las regiones más distantes, como la 6 u 8. Del mismo modo, la relación entre la inflación en cualquiera de las regiones 6 y 8 en el sur es mucho más cercana que aquéllas de las regiones norteñas con las sureñas. Una conclusión importante de esta inspección visual es que la geografía parece importar. Los choques regionales, o incluso las respuestas a los choques agregados en diferentes regiones, son probablemente dependientes de la ubicación de la región. Esta característica adicional de los datos será de utilidad para imponer parsimonia en el análisis empírico a continuación.

2 METODOLOGÍA

En esta sección, se presenta el marco econométrico utilizado para investigar las retroalimentaciones entre las tasas de inflación de n regiones en un país. Dos puntos importantes son considerados: el contraste de hipótesis y la identificación de choques estructurales. En cuanto a las pruebas de hipótesis, restricciones de agregación son impuestas en la forma reducida de un modelo de vectores autoregresivos (VAR) estándar y potencialmente grande de las tasas de inflación, y se propone un contraste para evaluar su significación formalmente. Estas restricciones promueven parsimonia mientras mantienen la retroalimentación dinámica de la tasa de inflación en una región determinada desde las tasas de inflación en el resto del país. En segundo lugar, se deriva una prueba para las restricciones de homogeneidad de largo plazo utilizando la forma de corrección de errores del modelo VAR. En el contexto de un modelo VAR de tasas de inflación, la homogeneidad de largo plazo puede ser entendida como el cumplimiento de la versión débil de la PPC, donde las diferencias entre las tasas de inflación son estacionarias.

Es importante destacar que los procedimientos de contraste de hipótesis propuestos han sido diseñados para ser robustos a las propiedades de estacionariedad de los datos. En particular, tanto la hipótesis agregación como la hipótesis de homogeneidad de largo plazo pueden ser contrastadas utilizando resultados estándares de la teoría de regresión clásica ya que los estadísticos en cuestión son asintóticamente chi-cuadrado, *independientemente* de si las series utilizadas son estacionarias o no. Los detalles técnicos se presentan en los anexos.

La forma de corrección de errores del VAR proporciona una formulación adecuada para inferir sobre las características de las tendencias comunes que conducen a los datos. En particular, se presta atención al caso donde la tasa de inflación en una región de referencia no responde a desequilibrios (desviaciones de la PPC relativa) en otras regiones. Así, se considera que un choque tiene efectos permanentes toda vez que afecte a la inflación en la región de referencia, y este choque puede ser fácilmente identificado utilizando la descomposición propuesta en Gonzalo y Ng (2001).

Hipótesis de agregación

El punto de partida es un modelo VAR (p) en forma reducida,

$$\mathbf{y}_t = \sum_{r=1}^p \mathbf{A}_r \mathbf{y}_{t-r} + \boldsymbol{\varepsilon}_t, \quad (1)$$

donde y_t es un vector de n variables endógenas cuyo i -ésimo elemento corresponde a la inflación de la región i en el periodo t , A_r ($r = 1, 2, \dots, p$) son matrices de coeficientes y ε_t es el vector de innovaciones estadísticas mutuamente correlacionadas. La matriz de covarianzas de ε_t es definida positiva.

Ha sido bien documentado que la utilidad de un modelo dinámico como (1) podría ser limitada en muestras finitas debido a la proliferación de los parámetros que deben estimarse. En efecto, cada rezago adicional implica la estimación de n^2 coeficientes, y estos pueden ser estimados muy imprecisamente con los tamaños de muestra con los que típicamente se dispone en aplicaciones. Por lo tanto, promover la parsimonia mediante la imposición de restricciones razonables sobre las matrices A_r podría mejorar el contenido inferencial de los procedimientos estadísticos basados en el sistema VAR. Éste es el propósito de las restricciones de agregación, donde ponderaciones dadas se utilizan para calcular promedios que mantienen efectos de retroalimentación entre las regiones.

Considere un agregado compuesto por las $n - 1$ variables en y_t distintas de $y_{i,t}$ ($i = 1, 2, \dots, n$),

$$x_{i,t} = \sum_{j=1}^n w_{ij} y_{j,t} \quad \text{donde} \quad \sum_{j=1}^n w_{ij} = 1 \quad \text{y} \quad w_{ii} = 0. \quad (2)$$

La normalización de que las ponderaciones suman uno es por conveniencia algebraica y no implica ninguna pérdida de generalidad. Se mantiene el supuesto de que las ponderaciones w_{ij} no son estimadas junto a A_r , de otro modo la linealidad en el modelo VAR puede perderse con la agregación. Esta situación corresponde tanto a ponderaciones no aleatorias como a ponderaciones estocásticas pero predeterminadas, es decir cuya determinación es independiente de ε_t (véase Winkelried y Saldarriaga, 2012).

La i -ésima ecuación del VAR sin restringir es

$$y_{i,t} = \sum_{r=1}^p a_{ii}(r) y_{i,t-r} + \sum_{r=1}^p \sum_{j \neq i}^n a_{ij}(r) y_{j,t-r} + \varepsilon_{i,t}, \quad (3)$$

donde $y_{i,t}$ es el i -ésimo elemento de y_t , $\varepsilon_{i,t}$ es el i -ésimo elemento de ε_t , y $a_{ij}(r)$ indica el (i, j) -ésimo elemento de A_r . En un modelo restringido alternativo, las retroalimentaciones hacia $y_{i,t}$ vienen de sus propios rezagos y de los rezagos del agregado:

$$y_{i,t} = \sum_{r=1}^p a_{ii}(r) y_{i,t-r} + \sum_{r=1}^p c_i(r) x_{i,t-r} + e_{i,t} = \sum_{r=1}^p a_{ii}(r) y_{i,t-r} + \sum_{r=1}^p \sum_{j \neq i}^n c_i(r) w_{ij} y_{j,t-r} + e_{i,t}. \quad (4)$$

Claramente, si $a_{ij}(r) = c_i(r) w_{ij}$ entonces el modelo restringido (4) equivale al modelo sin restricciones (3). Estas $p(n - 1)$ igualdades implican un total de $p(n - 1) - p = p(n - 2)$ restricciones de la forma

$$a_{ij}(r) - \left[\frac{w_{ij}}{w_{ik}} \right] a_{ik}(r) = 0 \quad \text{para} \quad j \neq k, k \neq i \quad \text{y} \quad r = 1, 2, \dots, p. \quad (5)$$

En consecuencia, las restricciones de agregación implican que los elementos de la i -ésima fila de A_r son proporcionales entre sí, y que el factor de proporcionalidad está dado por el ratio w_{ij}/w_{ik} . En otras palabras, $y_{j,t}$ y $y_{k,t}$ afectan el valor esperado de las futuras realizaciones de $y_{i,t}$ proporcionalmente a sus contribuciones al agregado (2).

El modelo sin restringir es obtenido al regresar $y_{i,t}$ sobre los p rezagos de y_t . Esto implica estimar pn coeficientes por ecuación y pn^2 en el VAR. Por otro lado, en el modelo restringido $y_{i,t}$ se regresa sobre sus p rezagos y los p rezagos del agregado $x_{i,t}$. Aquí, cada ecuación contiene $2p$ coeficientes y el VAR

restringido tiene $2pn$ coeficientes. Así, las restricciones de agregación reducen sustancialmente el número de coeficientes por estimar, incluso para valores moderados de n . Por ejemplo, si $p = 2$ y $n = 9$ se tiene $pn^2 = 162$ coeficientes en el modelo sin restringir, y sólo $2pn = 36$ en el restringido, un total de $np(n - 2) = 126$ restricciones.

Las restricciones de agregación pueden ser convenientemente interpretadas como restricciones de exclusión, y ésta es la base para las pruebas de hipótesis. Después de simples manipulaciones, la ecuación original (3) puede ser reescrita como

$$y_{i,t} = \sum_{r=1}^p a_{ii}(r)y_{i,t-r} + \sum_{r=1}^p c_i(r)x_{i,t-r} + \sum_{r=1}^p \sum_{j \neq i}^n \delta_{ij}(r)y_{j,t-r} + \varepsilon_{i,t}, \quad (6)$$

donde $\delta_{ij}(r) = a_{ij}(r) - c_i(r)w_{ij}$ para $r = 1, \dots, p$, $j = 1, 2, \dots, n$ y $j \neq i$. El modelo restringido tiene $\delta_{ij}(r) = 0$ para todo r y $j \neq i$, por lo que comprobar la hipótesis de agregación equivale a probar las restricciones de significación conjunta $H_0 : \delta_{ij}(r) = 0$ usando un estadístico de Wald estándar (ver anexo A). Nótese que H_0 tiene la interpretación tentativa de que una vez que se controla por $x_{i,t}$, sus componentes $y_{j,t}$ no presentan poder predictivo alguno sobre $y_{i,t}$.

Homogeneidad de largo plazo

El modelo restringido

$$y_{i,t} = \sum_{r=1}^p a_{ii}(r)y_{i,t-r} + \sum_{r=1}^p c_i(r)x_{i,t-r} + \varepsilon_{i,t}, \quad (7)$$

constituye una ecuación autoregresiva de rezagos distribuidos (ARDL) que describe la relación dinámica entre $y_{i,t}$ y $x_{i,t}$. El multiplicador de largo plazo $MLP_i = \sum_{r=1}^p c_i(r) / (1 - \sum_{r=1}^p a_{ii}(r))$ mide el efecto sobre $y_{i,t}$ de una desviación permanente de $x_{i,t}$ en una unidad con respecto a su nivel de largo plazo. La hipótesis de la PPC relativa entre $y_{i,t}$ y $x_{i,t}$ implica la condición de homogeneidad de largo plazo $MLP_i = 1$. Esto significa que la inflación de la región i ($y_{i,t}$) responde a una desviación permanente en el nivel de inflación en cualquier otro lugar ($x_{i,t}$) proporcionalmente, tal que el comportamiento de largo plazo de la inflación relativa ($y_{i,t} - x_{i,t}$) no se altere, es decir $y_{i,t} - x_{i,t}$ permanece estacionaria. Note que el cumplimiento de la condición de PPC relativa *por pares*, es decir $y_{i,t} - y_{j,t}$ es estacionaria para todo $i \neq j$, requiere que $MLP_i = 1$ para todas las regiones. Este punto se retoma más adelante.

Con el fin verificar si $MLP_i = 1$, la ecuación ARDL es escrita en su forma de corrección de errores

$$\Delta y_{i,t} = -\gamma_i(y_{i,t-1} - x_{i,t-1}) + \sum_{r=1}^{p-1} \pi_i^y(r)\Delta y_{i,t-r} + \sum_{r=1}^{p-1} \pi_i^x(r)\Delta x_{i,t-r} + \theta_i x_{i,t-1} + \varepsilon_{i,t}, \quad (8)$$

donde $\gamma_i = 1 - \sum_{r=1}^p a_{ii}(r) \geq 0$, $\pi_i^y(r) = -\sum_{s=r+1}^p a_{ii}(s)$, $\pi_i^x(r) = \sum_{s=r+1}^p c_i(s)$ y $\theta_i = \gamma_i(MLP_i - 1)$.

La hipótesis de homogeneidad de largo plazo puede ser escrita como $H_0 : \theta_i = 0$. Como se muestra en el anexo B, si (8) está correctamente especificada (es decir, $\varepsilon_{i,t}$ no se correlaciona con la historia de $x_{i,t}$), el estadístico t asociado con H_0 tiene la distribución asintótica normal usual, incluso cuando los datos no son estacionarios y la distribución límite de los estimadores mínimo cuadráticos de θ_i no es estándar. Ello provee una simple prueba de la hipótesis de la PPC relativa en la ecuación i .

Si $\gamma_i > 0$ entonces opera el denominado mecanismo de corrección de errores y así $y_{i,t}$ responde a las

desviaciones del agregado $x_{i,t}$ para restaurar el equilibrio de largo plazo. El coeficiente de “velocidad de ajuste” γ_i mide la proporción del desequilibrio $y_{i,t} - x_{i,t}$ que es corregido en cada periodo. Nótese que si $\gamma_i = 1$, entonces la transición hacia el equilibrio se da en un solo periodo. Por el contrario, $\gamma_i = 0$ corresponde al caso donde $y_{i,t}$ no responde al equilibrio en absoluto, $y_{i,t}$ es débilmente exógena (*long-run forcing*). Como se verá a continuación, este resultado puede ser explotado para identificar las tendencias que guían la dinámica del sistema formado por una colección de n ecuaciones como (8).

Identificación de los choques a partir de una región de referencia

Bajo la hipótesis de agregación y homogeneidad de largo plazo, el modelo VAR original (1) puede ser escrito como un modelo vectorial de corrección de errores (VECM)

$$\Delta \mathbf{y}_t = -\mathbf{\Pi} \mathbf{y}_{t-1} + \boldsymbol{\pi}_1 \Delta \mathbf{y}_{t-1} + \boldsymbol{\pi}_2 \Delta \mathbf{y}_{t-2} + \dots + \boldsymbol{\pi}_{p-1} \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t, \quad (9)$$

donde las matrices $\boldsymbol{\pi}$ incorporan las restricciones de agregación.² Es importante anotar que, por el teorema de representación de Granger (véase Hamilton, 1994, p. 581), un sistema de n variables guiadas por h tendencias estocásticas puede ser representada como un VECM como (9) con una matriz de largo plazo $\mathbf{\Pi}$ que tiene un rango igual a $n - h$. Ello implica que el sistema contiene $n - h$ relaciones de cointegración.

Sea \mathbf{W} la matriz $n \times n$ que recoge las ponderaciones w_{ij} (recuerde que $w_{ii} = 0$). Evaluando la definición en (2) dentro de (8), es simple verificar que la matriz $\mathbf{\Pi}$ toma la forma $\mathbf{\Pi} = \mathbf{\Gamma}_0 \mathbf{B}_0$ donde $\mathbf{\Gamma}_0$ es una matriz $n \times n$ diagonal con γ_i como su i -ésimo elemento diagonal, y $\mathbf{B}_0 = \mathbf{I}_n - \mathbf{W}$ es también una matriz de dimensión $n \times n$. Explícitamente,

$$\mathbf{\Pi} = \begin{bmatrix} \gamma_1 & 0 & 0 & \cdots & 0 \\ 0 & \gamma_2 & 0 & \cdots & 0 \\ 0 & 0 & \gamma_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \gamma_n \end{bmatrix} \begin{bmatrix} 1 & -w_{12} & -w_{13} & \cdots & -w_{1n} \\ -w_{21} & 1 & -w_{23} & \cdots & -w_{2n} \\ -w_{31} & -w_{32} & 1 & \cdots & -w_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -w_{n1} & -w_{n2} & -w_{n3} & \cdots & 1 \end{bmatrix}. \quad (10)$$

Sea $\mathbf{1}_m$ un vector suma (lleno de unos) de dimensión $m \times 1$. Luego, dado que por construcción $\mathbf{W} \mathbf{1}_n = \mathbf{1}_n$ (\mathbf{W} es *estocástica por filas*), resulta que $\mathbf{B}_0 \mathbf{1}_n = \mathbf{0}$, y así $\mathbf{1}_n$ es un vector propio de \mathbf{B}_0 asociado con un valor propio cero. Es fácil demostrar que la multiplicidad de los valores propios cero es uno, y así el rango de \mathbf{B}_0 es igual a $n - 1$. Si $\gamma_i > 0$ para al menos $n - 1$ ecuaciones, entonces el rango de $\mathbf{\Gamma}_0$ es $n - 1$ o n , lo que implica que $\mathbf{\Pi}$ es también de rango $n - 1$. De ello se desprende que en el sistema (9) debe haber una tendencia estocástica única y $n - 1$ relaciones de cointegración, $h = 1$.

La igualdad en (10) y su implicada tendencia común es una consecuencia directa de imponer $\theta_i = 0$ para todo i en (8): $y_{i,t}$ cointegra con $x_{i,t}$, y como ello ocurre para todo i , entonces $y_{i,t}$ cointegra con cada uno de los $n - 1$ elementos de $x_{i,t}$. Ello implica relaciones de cointegración por pares de la forma $y_{i,t} - y_{j,t} \sim I(0)$ para $i \neq j$, es decir el cumplimiento de la PPC relativa para cualquier par de regiones.³

² Defina a $\mathbf{\Pi}_r^x = \text{diag}(\pi_1^x(r), \pi_2^x(r), \dots, \pi_n^x(r))$ y $\mathbf{\Pi}_r^y = \text{diag}(\pi_1^y(r), \pi_2^y(r), \dots, \pi_n^y(r))$ como las matrices $n \times n$ diagonales que recogen los coeficientes asociados a los r -ésimos rezagos en (8). Luego, $\boldsymbol{\pi}_r = \mathbf{\Pi}_r^y + \mathbf{\Pi}_r^x \mathbf{W}$, donde \mathbf{W} es la matriz de ponderaciones definida más adelante. Nótese que la matriz $\boldsymbol{\pi}_r$ es de dimensión $n \times n$ y contiene sólo $2n$ parámetros libres.

³ Utilizando la descomposición (11) líneas abajo, el término de desequilibrio en (9) es $\mathbf{\Pi} \mathbf{y}_t = \mathbf{\Gamma} \mathbf{B} \mathbf{y}_t = \mathbf{\Gamma} \mathbf{z}_t$, donde \mathbf{z}_t es un vector de $(n - 1) \times 1$ cuyo i -ésimo elemento es el desvío $y_{i+1,t} - x_{i+1,t}$ ($i = 1, \dots, n - 1$). La descomposición (11) no es única y puede encontrarse una matriz no singular \mathbf{D} tal que $\mathbf{\Pi} \mathbf{y}_t = \mathbf{\Gamma} \mathbf{D} \mathbf{D}^{-1} \mathbf{B} \mathbf{y}_t = \mathbf{\Gamma}^\dagger \mathbf{w}_t$ donde el i -ésimo elemento de \mathbf{w}_t es el desvío $y_{i,t} - y_{1,t}$. Tal matriz es $\mathbf{D} = \mathbf{I} - \mathbf{W}_{22}$, ver ecuación (12). Así, el sistema originalmente expresado en términos de los desvíos $\mathbf{z}_t \sim I(0)$ es completamente equivalente a un sistema “rotado” expresado en términos de $\mathbf{w}_t \sim I(0)$, lo que conlleva inmediatamente a la conclusión que $w_{i,t} - w_{j,t} = y_{i,t} - y_{j,t} \sim I(0)$ para $i \neq j$.

Una representación VECM más familiar tendrá a $\mathbf{\Pi}$ descompuesta como el producto de una matriz con rango *columna* completo y una matriz con rango *fila* completo, ambas de rango $n - 1$. Puede llegarse fácilmente a dicha representación cuando uno de los coeficientes de ajuste γ_i es igual a cero. Sin ninguna pérdida de generalidad, suponga que $\gamma_1 = 0$, de modo que

$$\mathbf{\Pi} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \gamma_2 & 0 & \cdots & 0 \\ 0 & \gamma_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_n \end{bmatrix} \begin{bmatrix} -w_{21} & 1 & -w_{23} & \cdots & -w_{2n} \\ -w_{31} & -w_{32} & 1 & \cdots & -w_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -w_{n1} & -w_{n2} & -w_{n3} & \cdots & 1 \end{bmatrix}, \quad (11)$$

o compactamente $\mathbf{\Pi} = \mathbf{\Gamma}\mathbf{B}$, donde $\mathbf{\Gamma}$ es la matriz $n \times (n - 1)$ obtenida al remover la primera columna de $\mathbf{\Gamma}_0$, y \mathbf{B} es la matriz $(n - 1) \times n$ que resulta de remover la primera fila de \mathbf{B}_0 , ambas de rango de $n - 1$.

La representación alternativa (11) resulta muy conveniente para interpretar los choques que afectan al sistema y especialmente para identificar la tendencia común. La situación de que $\gamma_1 = 0$ y $\gamma_i > 0$ para $i > 1$ implica que $y_{1,t}$ es la única variable del sistema débilmente exógena (*long-run forcing*) y, siguiendo a Gonzalo y Granger (1995), esta misma variable puede ser utilizada para caracterizar la tendencia común. Intuitivamente, puede pensarse en cada $y_{i,t}$ como la acumulación del choque con efectos permanentes más un choque idiosincrático que desvía $y_{i,t}$ de la tendencia común. Dado que $y_{1,t}$ no se ajustará para restaurar el desequilibrio, entonces debe ser impulsada exclusivamente por el choque permanente, mientras $y_{i,t}$ para $i \neq 1$ responde al choque permanente (lo que mantiene $y_{i,t} - y_{1,t}$ estacionario) y también a las perturbaciones que transitoriamente apartan a $y_{i,t} - y_{1,t}$ de su expectativa de largo plazo.

Gonzalo y Ng (2001) proponen un método para expresar las perturbaciones del VECM $\boldsymbol{\varepsilon}_t$ como combinaciones lineales de h choques permanentes y $n - h$ choques transitorios, la “descomposición PT”. Así, el razonamiento anterior proporciona una manera sencilla de identificar el choque estructural que pueden atribuirse a la región 1: dado que sólo el $h = 1$ choque a la tendencia afecta $y_{1,t}$, entonces el choque permanente en una descomposición PT puede ser interpretado como proveniente de la región 1.

Formalmente, si \mathbf{u}_t denota al vector $n \times 1$ cuyo primer elemento es el choque permanente, $u_{1,t}$, y los $n - 1$ elementos restantes son choques transitorios, $\mathbf{u}_{2,t}$, Gonzalo y Ng muestran que $\mathbf{u}_{2,t} = \mathbf{B}\boldsymbol{\varepsilon}_t$, mientras $u_{1,t} = \mathbf{\Gamma}^*\boldsymbol{\varepsilon}_t$, donde $\mathbf{\Gamma}^*$ es un vector $1 \times n$ tal que $\mathbf{\Gamma}^*\mathbf{\Gamma} = \mathbf{0}$. De (11) queda claro que $\mathbf{\Gamma}^*$ debe ser proporcional al primer vector unitario en \mathbb{R}^n , es decir la primera fila de \mathbf{I}_n . De manera más compacta, $\mathbf{u}_t = \mathbf{G}\boldsymbol{\varepsilon}_t$, donde la primera fila de \mathbf{G} es $\mathbf{\Gamma}^*$ y su bloque inferior $(n - 1) \times n$ está dado por \mathbf{B} . Para tener una mejor comprensión sobre cómo los choques estructurales \mathbf{u}_t se relacionan con las innovaciones $\boldsymbol{\varepsilon}_t$, considere la igualdad $\boldsymbol{\varepsilon}_t = \mathbf{G}^{-1}\mathbf{u}_t$.⁴ Nótese que \mathbf{B} puede ser particionado como $\mathbf{B} = [-\mathbf{w}_{12} : \mathbf{I}_{n-1} - \mathbf{W}_{22}]$, donde \mathbf{w}_{12} y \mathbf{W}_{22} son bloques de \mathbf{W} . Dado que $(\mathbf{I}_{n-1} - \mathbf{W}_{22})\mathbf{1}_{n-1} = \mathbf{w}_{12}$, entonces

$$\mathbf{G}^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ -\mathbf{w}_{12} & \mathbf{I} - \mathbf{W}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ (\mathbf{I} - \mathbf{W}_{22})^{-1}\mathbf{w}_{12} & (\mathbf{I} - \mathbf{W}_{22})^{-1} \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{1}_{n-1} & (\mathbf{I} - \mathbf{W}_{22})^{-1} \end{bmatrix}. \quad (12)$$

La innovación de la primera ecuación equivale al choque permanente $\varepsilon_{1,t} = u_{1,t}$, y de esta manera es posible identificar cómo un choque en la región de referencia se propaga al resto del país. Nótese también que, en el impacto, el choque $u_{1,t} = 1$ tendrá un efecto proporcional en *todas* las innovaciones $\varepsilon_{i,t} = 1$

⁴ La práctica usual es ortogonalizar \mathbf{u}_t y considerar en cambio $\boldsymbol{\varepsilon}_t = \mathbf{G}^{-1}\mathbf{H}\boldsymbol{\eta}_t$, donde \mathbf{H} es el bloque inferior de la matriz triangular obtenida al aplicar una descomposición de Choleski a la matriz de covarianzas de \mathbf{u}_t , y $\boldsymbol{\eta}_t$ recoge los choques ortogonalizados. Ello no hace diferencia en la identificación del choque en la región de referencia. Al igual que en el caso de \mathbf{u}_t , el primer elemento de $\boldsymbol{\varepsilon}_t$ es afectado solamente por el primer elemento de $\boldsymbol{\eta}_t$, es decir el choque permanente ortogonalizado.

para $i = 1, 2, \dots, n$, mientras que un choque $u_{i,t} = 1$ cuando $i > 1$ afectará todas las innovaciones $(\varepsilon_{2,t}, \varepsilon_{3,t}, \dots, \varepsilon_{n,t})$ menos a $\varepsilon_{1,t}$.⁵

3 ANÁLISIS EMPÍRICO

A continuación se presentan los principales resultados del análisis empírico, usando tasas anualizadas de inflación mensual, $y_{i,t} = 1200 \log(P_{i,t}/P_{i,t-1})$. En primer lugar, se describen dos enfoques para medir las tasas de inflación del “resto del país”, uno basado en proximidad geográfica y el otro en influencia económica. Se encuentra evidencia que apoya la hipótesis de agregación bajo ambos enfoques. El VECM es entonces estimado y se encuentra que la PPC relativa se cumple en todo el país y, además, que la convergencia hacia una tendencia común es rápida. Adicionalmente, los resultados también indican que la influencia de Lima en el comportamiento de las inflaciones regionales, y así en la nacional, es fuerte.

Enfoques de agregación

Se consideran dos enfoques alternativos para construir el agregado $x_{i,t}$ lo que, dado que $x_t = Wy_t$, equivale a determinar la matriz de ponderaciones W . En ambos casos, los elementos de W están dados (son independientes de ε_t), y así se aplican las consideraciones teóricas sobre la linelidad después de la agregación discutida en la sección previa.

El primer enfoque corresponde a *agregación geográfica*. Éste se inspira en la noción básica de que la ubicación y la distancia importan, y en particular que la retroalimentación entre dos regiones será más fuerte cuanto más cercanas se encuentren. Existen diversas maneras de operacionalizar este concepto y se elige una que es particularmente simple (basada en relaciones de contigüidad entre las $n = 9$ regiones). Defina los valores

$$\tilde{w}_{ij} = \begin{cases} 0 & \text{si } i = j \text{ (las regiones no se consideran vecinas de sí mismas),} \\ 0 < a & \text{si la región } j \text{ e } i \text{ son contiguas,} \\ 0 \leq b < a & \text{si la región } j \text{ está localizada a una región de distancia de } i, \\ 0 \leq c < b & \text{de otro modo (} i \text{ y } j \text{ están separadas por al menos dos regiones),} \end{cases} \quad (13)$$

de forma que w_{ij} es simplemente una versión normalizada de \tilde{w}_{ij} , es decir $w_{ij} = \tilde{w}_{ij} / \sum_{j=1}^n \tilde{w}_{ij}$. Vale la pena mencionar que enfoques alternativos y más sofisticados de ponderación espacial reditúan resultados que son bastante similares a los reportados más adelante (disponibles bajo solicitud a los autores). Se establece arbitrariamente $a = 1$, $b = 0.5$ y $c = 0$, recordando que las principales conclusiones no son sensibles a esta elección.

El segundo enfoque es el de *agregación económica*. Aquí, la intensidad de la retroalimentación está determinada por ponderaciones que son proporcionales a la participación de cada región en el gasto nacional. El supuesto detrás de ello es que la importancia de la región i como un mercado de bienes intercambiados dentro del país es apropiadamente reflejada por su contribución relativa al gasto nacional. Bajo este enfoque, $x_{i,t}$ es una medida normalizada de la inflación nacional *sin la contribución de la*

⁵ Estos resultados pueden ser generalizados fácilmente. Suponga que hay dos regiones de referencia (dos tendencias comunes) tal que $\gamma_1 = \gamma_2 = 0$. En este caso, el rango de B_0 sigue siendo $n - 1$, pero $\text{rango}(\Pi_0) = \text{rango}(\Gamma_0) = n - 2$ por lo que en la descomposición $\Pi = \Gamma B$, Γ es la matriz $n \times (n - 2)$ obtenida al remover las dos primeras columnas de Γ_0 , y B es la matriz $(n - 2) \times n$ obtenida al remover las dos primeras filas de B_0 . Al examinar la estructura de la matriz G , se llega a la importante conclusión que las dos primeras innovaciones $(\varepsilon_{1,t}, \varepsilon_{2,t})$ son combinaciones lineales de choques permanentes $(u_{1,t}, u_{2,t})$, y ningún choque transitorio las afecta. Tras ortogonalizar u_t , o usar otro esquema de identificación, pueden encontrarse choques estructurales para las regiones 1 y 2. Para una discusión relacionada, véase Gonzalo y Granger (1995).

inflación en la región i . Sea Y_t la inflación nacional y sea α_j la ponderación que recibe la región j al componerla (estas ponderaciones son reportadas en el Gráfico 1). Por construcción, $\sum_{j=1}^n \alpha_j = 1$. Luego,

$$Y_t = \sum_{j=1}^n \alpha_j y_{j,t} = \alpha_i y_{i,t} + \sum_{j \neq i}^n \alpha_j y_{j,t} = \alpha_i y_{i,t} + (1 - \alpha_i) x_{i,t}, \quad (14)$$

así Y_t es una combinación lineal entre $y_{i,t}$ y $x_{i,t}$. Resulta que $x_{i,t} = \sum_{j \neq i}^n \alpha_j y_{j,t} / (1 - \alpha_i)$, por lo que

$$w_{ij} = \frac{\alpha_j}{1 - \alpha_i}. \quad (15)$$

En este caso, la matriz W no es simétrica. Una región grande (en términos económicos) ejerce una mayor influencia en una región pequeña que lo que una región pequeña ejerce sobre una más grande. De manera más formal, dado que $\alpha_i + \alpha_j < 1$ (lo cual siempre ocurre), resulta sencillo verificar que $w_{ij} < w_{ji}$ si y sólo si $\alpha_i > \alpha_j$. De este modo, la retroalimentación que la inflación en la región i recibe de la inflación en la región j es menor que la retroalimentación de la región i hacia la región j , porque la región i constituye, en términos de gasto, un mercado más grande que la región j .

Hipótesis de agregación

Al contrastar la hipótesis de agregación, una cuestión práctica de importancia es la determinación de número de rezagos p . Los resultados de la prueba LM de Breusch-Godfrey sobre las correlaciones seriales en los errores de la regresión (así como algunos criterios de información), sugieren utilizar $p \geq 5$. Esta sugerencia se mantiene tanto para modelos sin restringir (3) como para modelos restringidos (4). Así, se fija $p = 5$ ($p - 1 = 4$ rezagos en el VECM) por lo que bajo la hipótesis nula de agregación ($\delta_{ij}(r) = 0$ en la ecuación (6) para todo $i = 1, \dots, n$, $j \neq i$ y $r = 1, \dots, p$), el estadístico de Wald estándar se distribuye asintóticamente como χ^2 con $p(n - 2) = 35$ grados de libertad.

Como se puede apreciar en el Cuadro 1 (p. 91), la hipótesis de agregación no puede ser rechazada con niveles de significación convencionales en ningún caso, e independientemente del enfoque de agregación utilizado para calcular $x_{i,t}$ (geográfica o económica). Estos resultados pueden ser tomados como evidencia concluyente de que el modelo restringido, que utiliza promedios ponderados para resumir los efectos de retroalimentación desde el resto del país, es capaz de capturar las principales características de los datos. El siguiente paso, entonces, es investigar la dinámica de la forma restringida.

Homogeneidad de largo plazo

A diferencia de un VECM estándar, (9) constituye un sistema de regresiones aparentemente no relacionadas (SUR) con diferentes regresores. Por tanto, el método de estimación preferido es el de mínimos cuadrados generalizados (los resultados usando mínimos cuadrados ordinarios son cualitativamente similares).

Los resultados se reportan en el Cuadro 2 (p. 91), para $p = 5$. La primera columna muestra la prueba LM de Breusch-Godfrey y sugiere que cada ecuación dinámica está correctamente especificada. Este aspecto del modelo es relevante para realizar una inferencia válida sobre θ_i (véase el anexo B). La segunda y tercera columnas muestran las estimaciones puntuales y los p -values de γ_i y θ_i . Se puede observar que, en todas las ecuaciones, θ_i parece ser estadísticamente insignificante, lo que indica que la hipótesis de homogeneidad de largo plazo ($\theta_i = 0$ para todo i) no puede ser rechazada a niveles de significación convencionales. Por otro lado, las estimaciones de γ_i son estadísticamente distintas de cero, excepto en el caso de Lima. En consecuencia, el primer conjunto de resultados sugiere que el modelo VECM descrito

CUADRO 1. Contrastes de agregación

	Agregación geográfica		Agregación económica	
	Estadístico χ^2	<i>p</i> -value	Estadístico χ^2	<i>p</i> -value
1 Lima	31.654	0.630	32.295	0.599
2 Piura, Tumbes	35.446	0.447	38.924	0.297
3 Lambayeque, Cajamarca, Amazonas	27.838	0.800	31.513	0.637
4 La Libertad, Ancash	28.050	0.792	30.108	0.703
5 Ica, Ayacucho, Huancavelica	39.261	0.285	39.527	0.275
6 Arequipa, Moquegua, Tacna, Puno	28.221	0.785	31.129	0.656
7 Junín, Pasco, Huánuco	29.078	0.749	26.446	0.850
8 Cusco, Apurímac, Madre de Dios	37.764	0.344	36.229	0.411
9 Loreto, San Martín, Ucayali	35.291	0.454	38.136	0.329

NOTAS: Pruebas de Wald (asintóticamente χ^2) para $H_0 : \delta_{ij}(r) = 0$ en la ecuación (6), para todo $i = 1, \dots, n$, $j \neq i$ y $r = 1, \dots, p$, y para diferentes enfoques de agregación. El número de rezagos es $p = 5$, el mínimo valor que reedita residuos que no se encuentran serialmente correlacionados en todas las ecuaciones (tanto para los modelos sin restringir como para los restringidos). En todas las pruebas, el número de grados de libertad es $p(n - 2) = 35$.

anteriormente, con Lima como región de referencia, no puede ser rechazado por los datos. Estos hallazgos se mantienen para los enfoques de agregación geográfica y económica.

CUADRO 2. Estimaciones del VECM

	Prueba LM	γ_i	θ_i	γ_i	γ_i
	Breusch-Godfrey			($\theta_i = 0$)	($\theta_i = \gamma_1 = 0$)
<i>Agregación geográfica</i>					
1 Lima	0.074 (0.929)	0.548 (0.122)	0.274 (0.109)	0.399 (0.084)	–
2 Piura, Tumbes	2.036 (0.134)	1.128 (0.014)	0.194 (0.449)	1.056 (0.015)	1.059 (0.016)
3 Lambayeque, Cajamarca, Amazonas	0.474 (0.623)	0.769 (0.011)	0.224 (0.107)	0.743 (0.010)	0.787 (0.005)
4 La Libertad, Ancash	0.038 (0.963)	0.855 (0.034)	0.148 (0.545)	0.755 (0.028)	0.763 (0.022)
5 Ica, Ayacucho, Huancavelica	0.177 (0.838)	0.800 (0.058)	0.255 (0.108)	0.718 (0.041)	0.729 (0.032)
6 Arequipa, Moquegua, Tacna, Puno	1.478 (0.231)	0.997 (0.007)	0.197 (0.311)	0.982 (0.003)	1.000 (0.003)
7 Junín, Pasco, Huánuco	1.044 (0.354)	0.681 (0.023)	0.165 (0.211)	0.645 (0.020)	0.647 (0.025)
8 Cusco, Apurímac, Madre de Dios	2.046 (0.133)	0.692 (0.016)	0.057 (0.778)	0.635 (0.015)	0.625 (0.018)
9 Loreto, San Martín, Ucayali	3.390 (0.076)	0.806 (0.001)	0.245 (0.216)	0.704 (0.002)	0.647 (0.004)
<i>Agregación económica</i>					
1 Lima	0.084 (0.920)	0.125 (0.696)	0.178 (0.210)	0.071 (0.756)	–
2 Piura, Tumbes	1.318 (0.271)	1.044 (0.022)	0.043 (0.885)	0.969 (0.008)	0.960 (0.006)
3 Lambayeque, Cajamarca, Amazonas	0.238 (0.788)	0.900 (0.008)	0.099 (0.607)	0.890 (0.004)	0.873 (0.005)
4 La Libertad, Ancash	0.121 (0.886)	0.873 (0.008)	0.037 (0.876)	0.800 (0.005)	0.793 (0.005)
5 Ica, Ayacucho, Huancavelica	2.157 (0.119)	0.969 (0.008)	0.153 (0.382)	0.985 (0.002)	0.968 (0.002)
6 Arequipa, Moquegua, Tacna, Puno	1.242 (0.292)	0.848 (0.003)	0.112 (0.601)	0.842 (0.001)	0.834 (0.001)
7 Junín, Pasco, Huánuco	0.551 (0.577)	0.873 (0.001)	0.084 (0.510)	0.878 (0.000)	0.866 (0.000)
8 Cusco, Apurímac, Madre de Dios	1.351 (0.262)	0.787 (0.002)	0.021 (0.913)	0.750 (0.000)	0.749 (0.000)
9 Loreto, San Martín, Ucayali	1.707 (0.185)	0.806 (0.013)	0.166 (0.330)	0.790 (0.007)	0.787 (0.006)

NOTAS: Estimaciones SUR. La prueba Breusch-Godfrey es asintóticamente distribuida como una variable χ^2 con un grado de libertad bajo la hipótesis nula de ausencia de correlación serial de primer orden en los errores de regresión. Las cifras en paréntesis son *p*-values. Para los coeficientes estimados, los *p*-values son para la hipótesis de que el coeficiente es igual a cero.

La cuarta columna del Cuadro 2 muestra las estimaciones y *p-values* de γ_i después de imponer las restricciones de homogeneidad de largo plazo $\theta_i = 0$ para todo i . Nuevamente, todos los estimados excepto el de Lima son estadísticamente significativos. Por otra parte, las estimaciones puntuales de γ_i son considerablemente grandes para $i > 1$, en un rango entre 0.65 a 1.00 para la agregación geográfica y entre 0.75 a aproximadamente 1.00 para la agregación económica. Ello indica que las desviaciones de la PPC relativa se corrigen con bastante rapidez. La agregación geográfica implica un ajuste de algún modo más lento hacia la PPC, reflejando así la mayor influencia de la inflación de Lima (la tendencia común) en los agregados del “resto del país” bajo el enfoque de agregación económica. Todos estos resultados se confirman en la última columna del Cuadro 2, donde se presentan las estimaciones finales de γ_i , después de imponer las restricciones de homogeneidad de largo plazo y el hecho que Lima hace las veces de región de referencia ($\gamma_1 = 0$).⁶

Dinámica regional y la importancia de Lima

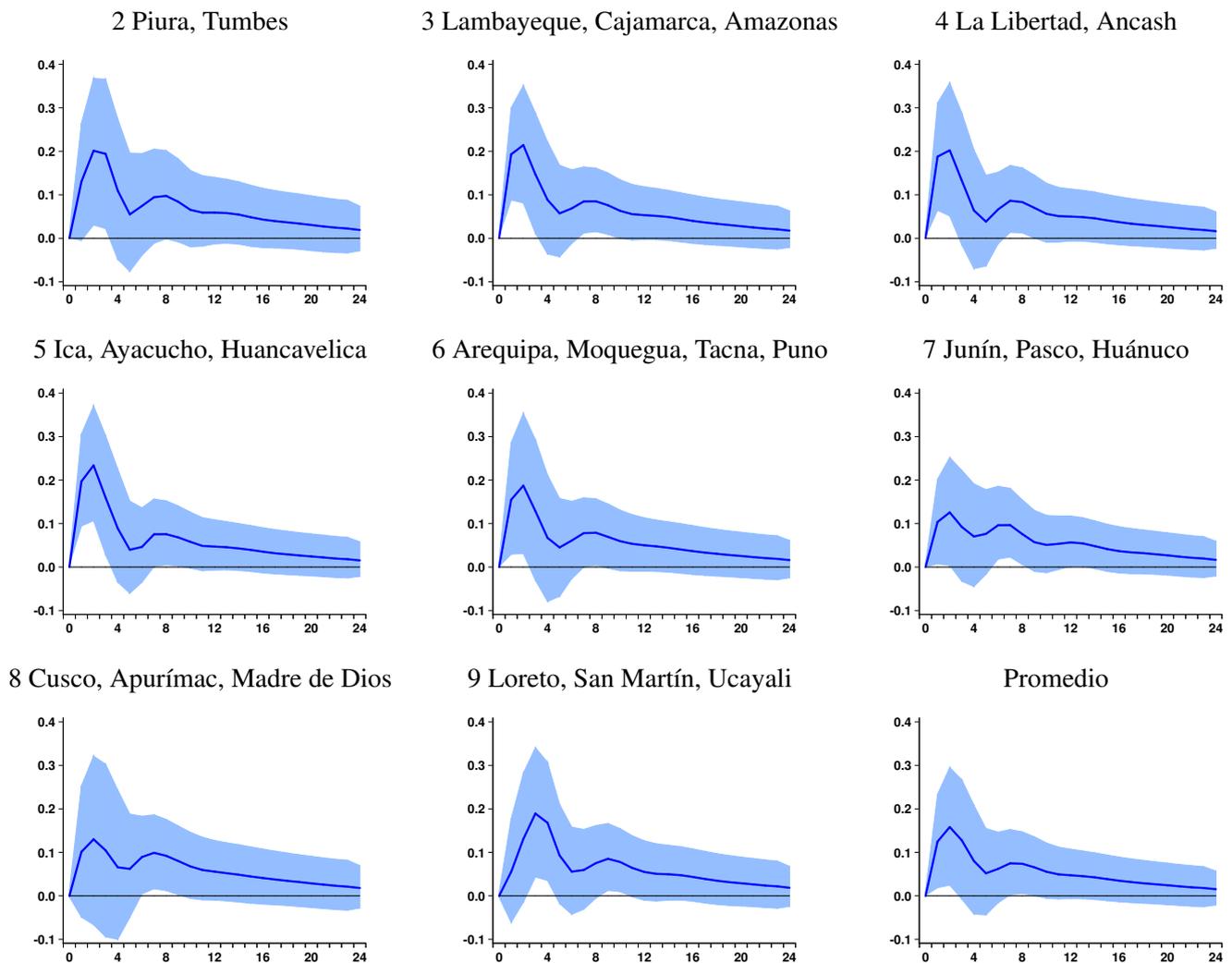
Las estimaciones revelan una rápida convergencia de las inflaciones regionales hacia sus propios agregados de inflación externa, $y_{i,t} - x_{i,t}$. Dado que este fenómeno ocurre en cada región del país, podría esperarse una rápida reversión hacia la tendencia común. El Gráfico 3 (p. 93) muestra las respuestas de las inflaciones regionales relativas a Lima, ante un choque unitario en la inflación de Lima (el choque permanente del sistema), bajo el enfoque de agregación geográfica (los resultados bajo el enfoque de agregación económica son muy similares y están disponibles a solicitud). A modo de resumen, el último panel del gráfico muestra la respuesta del promedio de las inflaciones de las $n = 9$ regiones en el país.

Debido a la dinámica intrínseca causada tanto por los rezagos del VECM como por las fricciones geográficas capturadas en la matriz W , la convergencia de $y_{i,t}$ hacia la tendencia común $y_{1,t}$ no es tan inmediata como aquella hacia $x_{i,t}$. Las funciones de impulso-respuesta presentan “jorobas”, sugiriendo una sobre-reacción inicial de $y_{i,t}$ al choque, que a la larga se desvanece. No obstante, la convergencia es rápida: la desviación máxima equivale a menos de 0.4 por ciento en términos anuales, y además ninguna respuesta es estadísticamente significativa después de 12 meses de la ocurrencia del choque (varias de ellas no son significativas después de 6 meses). Por lo tanto, puede concluirse que un choque a la inflación tendencial se propagará completamente hacia las inflaciones regionales individuales en menos de un año.

Por otro lado, como se discutió en la sección 1, las inflaciones regionales están afectadas también por choques idiosincrásicos que pueden inducir a desviaciones de corto plazo respecto a la tendencia. El Gráfico 4 (p. 94) muestra las respuestas de las inflaciones relativas a Lima ante un choque en la región i que induce a una desviación en el impacto de uno por ciento, bajo el enfoque de agregación geográfica. Se observa que los efectos de los choques son de corta duración: para todos los casos, después de tres meses del choque, la desviación con respecto a la inflación de Lima es de menos de 0.2 por ciento. Dada la naturaleza transitoria de estos choques, pocas respuestas aparecen como estadísticamente significativas después de 12 meses del choque y ninguna es significativa después de 18 meses.

Para evaluar la importancia del choque a la tendencia común, se calcula la descomposición de varianza de los errores de predicción en el VECM. El Cuadro 3 (p. 95) muestra la contribución del choque permanente para diferentes horizontes de predicción. El largo plazo $h \rightarrow \infty$ corresponde a la varianza no condicional y vale la pena enfocarse en estas cifras primero. La importancia de la tendencia común no puede exagerarse: bajo el enfoque de agregación geográfica, los choques permanentes explican al

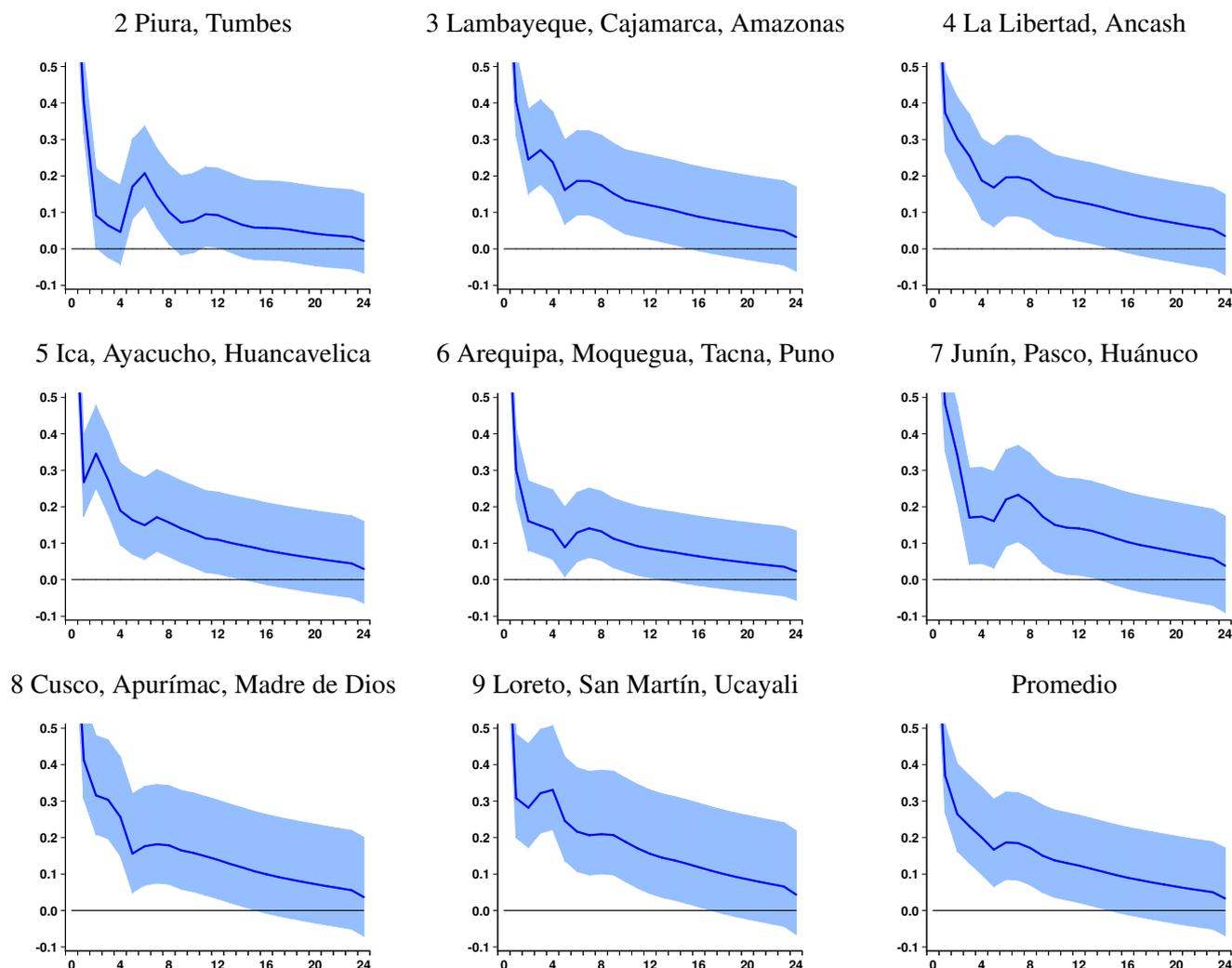
⁶ Bajo agregación geográfica, el estadístico de Wald de la hipótesis conjunta $\theta_i = 0$ para todo i es $\omega_1 = 10.32$, por debajo del valor crítico 90%, $\chi_8^2 = 14.68$. Asimismo, el estadístico de Wald para la hipótesis conjunta de que $\gamma_i = 0$ para $i > 1$ es $\omega_2 = 55.79$, muy por encima del valor crítico 99%, $\chi_8^2 = 20.09$. Bajo el esquema agregación económica, estos estadísticos son aún más categóricos: $\omega_1 = 4.77$ y $\omega_2 = 78.69$.

GRÁFICO 3. Respuesta de las inflaciones regionales relativas a la de Lima ante un choque en Lima

NOTAS: Resultados bajo el esquema de agregación geográfica. Cada panel presenta la respuesta de $y_{i,t} - y_{1,t}$ (en términos porcentuales anualizados) a un choque unitario en Lima, identificado como el choque permanente de la descomposición PT. El eje horizontal representa meses después de ocurrido el choque en el periodo 0. Las áreas sombreadas representan intervalos *bootstrap* al 90 por ciento de confianza. El eje vertical es el mismo para todos los gráficos.

menos 69.1 por ciento de la varianza no condicional (en la región 8) y hasta 83.7 por ciento (en la región 7). La contribución a la inflación promedio resulta ser un muy relevante 89.7 por ciento. Como era de esperarse, bajo el enfoque de agregación económica, la importancia de la inflación de Lima se incrementa aún más: el choque contribuye al menos con 74.6 por ciento de la varianza no condicional (nuevamente en la región 8), y con más de 95 por ciento para la tasa de inflación promedio. La evolución de la descomposición de la varianza implica que la influencia de un choque de la tendencia común es notoria un año en adelante, mientras que después de dos años las cifras son cercanas a los estadísticos no condicionales. Estos hallazgos reflejan que lo que ocurre en Lima sirve como punto de referencia en las decisiones detrás de la fijación de precios en todo el país, en un horizonte relevante para la política monetaria, digamos entre uno y dos años.⁷

⁷ El caso de la región 8 merece especial atención ya que se trata de la región donde los choques transitorios parecen ser más importantes. Esta región incluye a la ciudad del Cusco, la cual es de lejos el mayor destino turístico del país, y así su dinámica en la fijación los precios (especialmente en servicios) probablemente se desacople del comportamiento del resto de economías regionales, Lima entre ellas.

GRÁFICO 4. Respuesta de las inflaciones regionales relativas a la de Lima ante choques regionales

NOTAS: Resultados bajo el esquema de agregación geográfica. Cada gráfico presenta la respuesta de $y_{i,t} - y_{1,t}$ (en términos porcentuales anualizados) a un choque en la región i tal que $y_{i,0} - y_{1,0} = 1$. El eje horizontal representa meses después de ocurrido el choque en el periodo 0. Las áreas sombreadas representan intervalos *bootstrap* al 90 por ciento de confianza. El eje vertical es el mismo para todos los gráficos.

4 OBSERVACIONES FINALES

En este trabajo se ha estimado y examinado un modelo multivariado dinámico con el objetivo de cuantificar las retroalimentaciones entre las tasas de inflación regionales en el Perú. En particular, se ha evaluado la importancia de la inflación de Lima para el resto del país. Las estimaciones sugieren que las principales características de los datos pueden ser bien descritas por un VECM con dos peculiaridades. Primero, el modelo es muy parsimonioso, ya que las retroalimentaciones pueden ser resumidas a través de agregados de las regiones vecinas (tanto en un sentido geográfico como en uno económico) en lugar de considerar a las inflaciones individuales. Segundo, las estimaciones proveen evidencia de que la PPC relativa se cumple entre los pares de inflaciones regionales, y de que la única variable identificable como débilmente exógena es la inflación de Lima.

Estos hallazgos permiten concluir que la inflación de Lima ha servido como una tendencia para el resto de inflaciones regionales. Adicionalmente, la desviación típica respecto a esta tendencia es moderada y de

CUADRO 3. Contribución de un choque en Lima (permanente) a la varianza del error de predicción

	$h = 0$	$h = 1$	$h = 6$	$h = 12$	$h = 18$	$h = 24$	$h \rightarrow \infty$
<i>Agregación geográfica</i>							
2 Piura, Tumbes	51.8	55.5	60.3	65.9	70.3	73.7	77.2
3 Lambayeque, Cajamarca, Amazonas	56.5	63.2	68.8	73.9	77.6	80.6	83.3
4 La Libertad, Ancash	49.3	57.3	64.3	70.6	75.1	78.4	80.9
5 Ica, Ayacucho, Huancavelica	56.5	61.8	67.8	74.0	78.2	81.3	83.3
6 Arequipa, Moquegua, Tacna, Puno	44.4	50.0	57.9	66.4	72.0	76.0	78.2
7 Junín, Pasco, Huánuco	53.7	56.7	66.3	73.9	78.8	82.2	83.7
8 Cusco, Apurímac, Madre de Dios	18.2	25.1	39.9	53.2	61.2	66.8	69.1
9 Loreto, San Martín, Ucayali	40.0	47.3	60.3	68.7	74.1	77.9	80.1
Promedio	76.1	76.3	79.0	82.7	85.5	87.6	89.7
<i>Agregación económica</i>							
2 Piura, Tumbes	50.5	57.3	64.4	71.4	76.1	79.4	81.2
3 Lambayeque, Cajamarca, Amazonas	56.4	66.2	74.8	81.1	84.8	87.3	88.3
4 La Libertad, Ancash	49.7	59.3	68.3	75.7	80.3	83.3	84.6
5 Ica, Ayacucho, Huancavelica	58.5	68.4	77.1	83.3	86.8	89.1	89.8
6 Arequipa, Moquegua, Tacna, Puno	46.2	53.8	65.1	73.6	78.9	82.3	83.6
7 Junín, Pasco, Huánuco	58.2	62.7	77.6	84.9	88.7	91.0	91.4
8 Cusco, Apurímac, Madre de Dios	21.0	28.0	46.0	59.5	67.6	73.0	74.6
9 Loreto, San Martín, Ucayali	41.3	50.8	67.7	77.0	82.2	85.4	86.5
Promedio	78.2	82.2	88.0	91.8	93.7	94.9	95.3

NOTAS: El choque a la inflación de Lima está identificado como se discutió en la sección 2.0.3. Las figuras muestran el porcentaje de la varianza del error de predicción que es atribuible a este choque, para varios horizontes de predicción h . El caso $h \rightarrow \infty$ corresponde a la contribución de la variable no condicional del error de predicción.

corta duración. Dado que estas estimaciones provienen de una muestra donde el BCRP ha monitoreado de forma activa y explícita la inflación de Lima, se puede inferir que dichas prácticas han contribuido a estabilizar la inflación en todo el país.

De cara al futuro, se espera que la importancia económica de Lima en componer el gasto agregado nacional se mantenga alta en el mediano plazo. Así, el mercado puede seguir percibiendo a la inflación de Lima como una medida de la inflación nacional “libre de ruido”, útil para las decisiones de fijación de precios. Asimismo, es poco probable que la adopción de una meta de inflación nacional debilite la fuerte retroalimentación de la inflación de Lima a la inflación en el resto del país. Por lo tanto, en el evento que el BCRP opte por monitorear la inflación nacional, los resultados de este estudio indican que el funcionamiento de la política monetaria en el Perú (especialmente el rezago de la política monetaria) no debería verse afectado significativamente.

De acuerdo con hallazgos en Monge y Winkelried (2009), la mayor proporción de alimentos en la canasta del IPC en las provincias podría estar guiando los resultados hacia la rápida convergencia de la inflación, ya que los precios de alimentos están sujetos a choques de oferta que, a pesar de su tamaño, se revierten rápidamente. Consecuentemente, una extensión natural de este trabajo es repetir el análisis utilizando IPC regionales de cierta forma armonizados, por ejemplo utilizando la mismas ponderaciones que en el IPC de Lima. Alternativamente, el análisis puede realizarse usando medidas de inflación *subyacente*, excluyendo elementos volátiles tales como productos alimenticios y combustibles. A pesar de que estas extensiones seguramente mejorarán nuestra comprensión de la compleja transmisión de choques regionales en el país, se espera que los resultados cualitativos de este documento se mantengan.

ANEXO TÉCNICO

Siguiendo desarrollos teóricos en Phillips (1986), Park y Phillips (1988) y Sims y otros (1990), este anexo muestra que tanto la hipótesis de agregación como la hipótesis de homogeneidad de largo plazo pueden ser contrastadas mediante resultados estándares de la teoría de regresión clásica. En particular, se muestra que las pruebas de Wald involucradas son asintóticamente χ^2 cuando los datos no son estacionarios.

A HIPÓTESIS DE AGREGACIÓN

El modelo VAR $\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \mathbf{A}_2\mathbf{y}_{t-2} + \dots + \mathbf{A}_p\mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t$ puede ser escrito como

$$\mathbf{y}_t = \mathbf{C}_1\Delta\mathbf{y}_{t-1} + \mathbf{C}_2\Delta\mathbf{y}_{t-2} + \dots + \mathbf{C}_{p-1}\Delta\mathbf{y}_{t-p+1} + \mathbf{H}\mathbf{y}_{t-1} + \boldsymbol{\varepsilon}_t, \quad (\text{A1})$$

donde $\mathbf{H} = \mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_p$, $\mathbf{C}_1 = \mathbf{A}_1 - \mathbf{H}$ y $\mathbf{C}_r = \mathbf{C}_{r-1} + \mathbf{A}_r$ para $r = 2, \dots, p-1$.

Si \mathbf{y}_t es estacionario, los estimadores mínimo cuadráticos de las matrices \mathbf{C} y de \mathbf{H} serán asintóticamente normales bajo condiciones generales. Por otro lado, siguiendo a Sims y otros (1990) y Hamilton (1994, sección 18.2), cuando \mathbf{y}_t no es estacionario, los estimadores de las matrices \mathbf{C} son consistentes y asintóticamente normales, mientras que el estimador de \mathbf{H} es superconsistente y tiene una distribución límite que no es estándar. De manera más formal, se encuentra que para $r = 1, 2, \dots, p-1$

$$\sqrt{T}\text{vec}(\hat{\mathbf{C}}_r - \mathbf{C}_r) \xrightarrow{d} N(\mathbf{0}, \mathbf{V}_r) \quad \text{mientras que} \quad \sqrt{T}\text{vec}(\hat{\mathbf{H}} - \mathbf{H}) = o_p(1), \quad (\text{A2})$$

donde $\text{vec}(\mathbf{A})$ es el vector que resulta de concatenar las columnas de la matriz \mathbf{A} , y \mathbf{V}_r denota una matriz de covarianzas definida positiva. Los resultados en (A2) indican que las pruebas de hipótesis que involucran combinaciones lineales de las \mathbf{C} matrices y \mathbf{H} serán dominadas asintóticamente por los estimadores con una tasa de convergencia más lenta, es decir aquéllos de las matrices \mathbf{C} . En particular, cualquier prueba que involucre combinaciones lineales de las matrices \mathbf{A} diferentes a la suma \mathbf{H} presentará una distribución asintótica estándar.

Recuerde que para $r = 1, 2, \dots, p$ las restricciones de agregación tienen la forma $a_{ij}(r) - (w_{ij}/w_{ik})a_{ik}(r)$ para $j \neq k$ y $k \neq i$. De manera más compacta, $\mathbf{R}_r\text{vec}(\mathbf{A}_r) = \mathbf{0}$, donde \mathbf{R}_r es una matriz $p(n-2) \times n^2$ cuya fila típica contiene un 1, el ratio $-w_{ij}/w_{ik}$ y $n^2 - 2$ ceros. En términos de los coeficientes del modelo reparametrizado (A1), las restricciones de agregación son (considere $\mathbf{C}_p = \mathbf{0}$)

$$\mathbf{R}_r\text{vec}(\mathbf{A}_r) \equiv \mathbf{R}_r\text{vec}(\mathbf{C}_r - \mathbf{C}_{r-1}) = \mathbf{0} \quad \text{para} \quad r = 2, \dots, p, \quad (\text{A3})$$

$$\mathbf{R}_1\text{vec}(\mathbf{A}_1) \equiv \mathbf{R}_1\text{vec}(\mathbf{C}_1 + \mathbf{H}) = \mathbf{0}.$$

Consecuentemente, utilizando los resultados en (A2), se obtiene que bajo la hipótesis nula en (A3) (considere $\hat{\mathbf{C}}_p = \mathbf{0}$),

$$\sqrt{T}\mathbf{R}_r\text{vec}(\hat{\mathbf{C}}_r - \hat{\mathbf{C}}_{r-1}) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Omega}_r) \quad \text{para} \quad r = 2, \dots, p, \quad (\text{A4})$$

$$\sqrt{T}\mathbf{R}_1\text{vec}(\hat{\mathbf{C}}_1 + \hat{\mathbf{H}}) = \sqrt{T}\mathbf{R}_1\text{vec}(\hat{\mathbf{C}}_1) + o_p(1) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Omega}_1),$$

donde $\boldsymbol{\Omega}_r$ denota matrices de covarianzas definidas positivas para $r = 1, 2, \dots, p$. Se deduce que, independientemente de las propiedades de estacionariedad de los datos, los estimadores que participan en la prueba de la hipótesis de agregación son asintóticamente normales y, por lo tanto, las pruebas de Wald correspondientes serán asintóticamente χ^2 .

B HOMOGENEIDAD DE LARGO PLAZO

Considere una versión estilizada de la ecuación (8) con $p = 1$,

$$\Delta y_t = -\gamma(y_{t-1} - x_{t-1}) + \theta x_{t-1} + \varepsilon_t, \quad (\text{B1})$$

donde $\Delta x_t = v_t$ y ε_t son procesos estacionarios con media cero y varianzas σ_v^2 y σ_ε^2 , respectivamente. Se asume que v_t y ε_t no están correlacionados en ninguno de sus adelantos y rezagos, lo cual esencialmente implica que la ecuación (B1) está correctamente especificada. Los argumentos a continuación son válidos también para ecuaciones aumentadas que incluyan rezagos de Δx_t y Δy_t , con el propósito de garantizar la ausencia de correlación entre v_t y ε_t .

El objetivo es contrastar $H_0 : \theta = 0$ en (B1). Por brevedad, defina $z_t = y_t - x_t$, que es estacionario bajo H_0 . No resulta difícil verificar que, dado que $z_t \sim I(0)$ y que $x_t \sim I(1)$,

$$\frac{1}{T} \frac{\sum_t x_{t-1} z_{t-1} \sum_t z_{t-1} \varepsilon_t}{\sum_t (z_{t-1})^2} = \frac{1}{T} \frac{O_p(T) O_p(\sqrt{T})}{O_p(T)} = o_p(1) \quad \text{y} \quad \frac{1}{T^2} \frac{(\sum_t x_{t-1} z_{t-1})^2}{\sum_t (z_{t-1})^2} = \frac{1}{T^2} \frac{O_p(T^2)}{O_p(T)} = o_p(1). \quad (\text{B2})$$

Por otro lado, sean $W_\varepsilon(\cdot)$ y $W_v(\cdot)$ dos procesos de Weiner estándares, definidos en $C(0, 1)$ y asociados con las series estandarizadas $\varepsilon_t/\sigma_\varepsilon$ y v_t/σ_v , respectivamente. Estos procesos no están correlacionados y, por lo tanto, dada la normalidad de los incrementos, son independientes. Un resultado conocido para procesos integrados (véase Hamilton, 1994, sección 17.5) es que

$$\frac{1}{T^2} \sum_t (x_{t-1})^2 \xrightarrow{d} \sigma_v^2 \int_0^1 W_v(r)^2 dr \equiv \mathcal{D}_2, \quad (\text{B3})$$

mientras que Phillips (1986, p. 327) y Park y Phillips (1988, Lema 5.1) muestran que

$$\frac{1}{T} \sum_t x_{t-1} \varepsilon_t \xrightarrow{d} \sigma_v \sigma_\varepsilon \int_0^1 W_v(r) dW_\varepsilon(r) = N(0, \sigma_\varepsilon^2 \mathcal{D}_2) \equiv \mathcal{D}_1. \quad (\text{B4})$$

La distribución asintótica \mathcal{D}_1 es una mixtura de normales. Puede pensarse en esta distribución como una distribución normal con varianza proporcional a \mathcal{D}_2 que es en sí misma una realización aleatoria en el espacio de escalares positivos, en este caso funcionales cuadráticos de un proceso de Weiner.

Finalmente, sea s^2 el estimador usual de máxima verosimilitud de la varianza del error de la regresión, es decir el promedio de la suma del cuadrado de los residuos en la regresión (B1), que es consistente para σ_ε^2 . Al juntar los resultados en (B2), (B3) y (B4), el estadístico t para la prueba $H_0 : \theta = 0$ satisface, bajo esta hipótesis nula,

$$\tau = \frac{\sum_t x_{t-1} \varepsilon_t - \sum_t x_{t-1} z_{t-1} \sum_t z_{t-1} \varepsilon_t / \sum_t (z_{t-1})^2}{\sqrt{s^2 \sum_t (x_{t-1})^2 - s^2 (\sum_t x_{t-1} z_{t-1})^2 / \sum_t (z_{t-1})^2}} = \frac{\sum_t x_{t-1} \varepsilon_t / T}{\sqrt{s^2 \sum_t (x_{t-1})^2 / T^2}} + o_p(1) \xrightarrow{d} \frac{\mathcal{D}_1}{\sqrt{\sigma_\varepsilon^2 \mathcal{D}_2}} \equiv N(0, 1), \quad (\text{B5})$$

que se deriva de los teoremas del mapeo continuo y de Cramér. Por lo tanto, la distribución asintótica de τ es normal estándar, incluso cuando el estimador mínimo cuadrático de θ no es asintóticamente normal (la distribución límite de $T\hat{\theta}$ es $\mathcal{D}_1/\mathcal{D}_2$). Argumentos similares muestran que las pruebas de Wald para el contraste de restricciones generales de los coeficientes θ en varias de las ecuaciones del VECM (por ejemplo, τ^2 para el caso univariado) son asintóticamente χ^2 .

REFERENCIAS

- Armas, A., L. Vallejos y M. Vega (2010), "Measurement of price indices used by the Central Bank of Peru", en Bank for International Settlements (ed.), *Monetary Policy and the Measurement of Inflation: Prices, Wages and Expectations*, BIS Papers chapters 49, 259-283.
- Beck, G. W., K. Hubrich y M. Marcellino (2009), "Regional inflation dynamics within and across Euro area countries and a comparison with the United States", *Economic Policy*, 24(57), 141-184.
- Buseti, F., S. Fabiani y A. C. Harvey (2006), "Convergence of prices and rates of inflation", *Oxford Bulletin of Economics and Statistics*, 68(S1), 863-877.
- Cecchetti, S. G., C. M. Nelson y R. J. Sonora (2002), "Price index convergence among United States cities", *International Economic Review*, 43(4), 1081-1099.
- Gonzales de Olarte, E. (2003), *Regiones Integradas: Ley de Incentivos para la Integración y Conformación de Regiones. Lineamientos Económicos y Políticos*, Fondo Editorial del Congreso del Perú.
- Gonzalo, J. y C. W. J. Granger (1995), "Estimation of common long-memory components in cointegrated systems", *Journal of Business and Economic Statistics*, 13(1), 27-35.
- Gonzalo, J. y S. Ng (2001), "A systematic framework for analyzing the dynamic effects of permanent and transitory shocks", *Journal of Economic Dynamics and Control*, 25(10), 1527-1546.
- Hamilton, J. D. (1994), *Time Series Analysis*, Princeton University Press.
- Monge, A. y D. Winkelried (2009), "PPC y convergencia de precios: Un análisis para las ciudades del Perú", Banco de la República, *Ensayos de Política Económica*, 27(58), 56-105.
- Park, J. Y. y P. C. B. Phillips (1988), "Statistical inference in regressions with integrated processes: Part 1", *Econometric Theory*, 4(3), 468-497.
- Parsley, D. C. y S. J. Wei (1996), "Convergence to the law of one price without trade barriers or currency fluctuations", *Quarterly Journal of Economics*, 111(4), 1211-1236.
- Pesaran, M. H., T. Schuermann y S. M. Weiner (2004), "Modeling regional interdependencies using a global error-correcting macroeconometric model", *Journal of Business and Economic Statistics*, 22(2), 129-162.
- Phillips, P. C. B. (1986), "Understanding spurious regressions in econometrics", *Journal of Econometrics*, 33(3), 311-340.
- Rogers, J. H. (2007), "Monetary union, price level convergence, and inflation: How close is Europe to the USA?", *Journal of Monetary Economics*, 54(3), 785-796.
- Rossini, R. y M. Vega (2008), "The monetary policy transmission mechanism under financial dollarisation: The case of Peru 1996 - 2006", en Bank for International Settlements (ed.), *Transmission Mechanisms for Monetary Policy in Emerging Market Economies*, BIS Papers chapters 35, 395-412.
- Sims, C. A., J. H. Stock y M. W. Watson (1990), "Inference in linear time series models with some unit roots", *Econometrica*, 58(1), 113-144.
- Winkelried, D. y M. A. Saldarriaga (2012), "Trade linkages and growth in Latin America: A time-varying SVAR approach", Banco Central de Reserva del Perú, Working Paper 2011-012.