

CAPÍTULO III

Diferencias provinciales en la evolución del índice de precios al consumo

Antonio Montañés*

Este documento aborda el análisis de la evolución de los precios en España utilizando datos desagregados tanto geográfica (52 provincias) como temporalmente (período 2002:1-2023:9). Para ello, se ha utilizado una metodología que permite contrastar tanto la presencia de un único patrón de comportamiento común (convergencia) como la presencia de varios clubes de convergencia. Los resultados muestran que, para el total del índice de precios al consumo, existen hasta cinco patrones de comportamiento significativamente diferentes. Este resultado se mantiene si se descompone este índice en sus componentes principales, a excepción de los de bebidas alcohólicas y tabaco (G02) y de transporte (G07).

Palabras clave: precios, disparidades provinciales, convergencia, factores comunes, Phillips-Sul.

* Agradezco la invitación cursada por D. Peña y M. P. Poncela para participar en este evento. La actual versión se ha visto beneficiada por los comentarios de un evaluador anónimo.

1. INTRODUCCIÓN

Uno de los avances más destacados en el campo del análisis cuantitativo es la amplia disponibilidad de datos a la que los investigadores tienen acceso en la actualidad. El aumento en la capacidad de almacenamiento, combinado con el mayor rendimiento de las computadoras, ha llevado a los investigadores a adaptarse a estas nuevas circunstancias. En muchos casos, esto ha dado lugar al desarrollo de nuevas técnicas que permiten el manejo de datos masivos. Es crucial destacar las mejoras significativas en los métodos de *machine learning*, las capacidades mejoradas de las técnicas de regresión en paralelo y los algoritmos de optimización más rápidos y eficientes en comparación con los enfoques tradicionales, entre otras ventajas.

Sin embargo, también resulta necesario reconocer que esta disponibilidad masiva de datos también ha planteado ciertos desafíos. Por ejemplo, los investigadores han tenido que desarrollar métodos para seleccionar de manera adecuada las variables a utilizar. Entre estos métodos, se deben mencionar en particular el LASSO o *Adaptative Lasso* (Zhou, 2006; Chatterjee *et al.*, 2015), así como el uso de factores comunes para resumir la información contenida en un conjunto amplio de variables, como lo hacen Stock y Watson (2002) en el ámbito de la predicción.

Esta abundancia de datos y las nuevas herramientas econométricas han beneficiado significativamente las investigaciones en áreas como el análisis financiero o la medicina, especialmente en epidemiología. No obstante, posiblemente el campo que ha aprovechado mejor estas bases de datos masivas sea la economía. En particular, los estudios de convergencia han experimentado un impulso significativo. Los trabajos pioneros de Baumol (1986) o Barro y Sala-i-Marti (1992) buscaban analizar la posible existencia de procesos de *catching-up*, mediante los cuales el producto interno bruto per cápita de una región tiende a acercarse con el tiempo al de otra región más próspera, reduciendo las distancias inicialmente existentes y, en algunos casos, llegando a converger con este último.

Recientemente, Phillips y Sul (2007) han desarrollado métodos que permiten estudiar la hipótesis nula de convergencia. Este método es muy flexible en cuanto a las características temporales de las variables y, lo que es crucial en nuestro caso, se adapta de manera eficiente a la existencia de volúmenes de información elevados, facilitando el estudio de un patrón de comportamiento común para datos muy desagregados tanto geográfica como temporalmente.

El objetivo de este trabajo es utilizar precisamente esta herramienta para analizar la evolución de los precios en España, abarcando un intervalo temporal extenso y, sobre todo, una desagregación territorial tan amplia como lo permitan los datos. En nuestro caso, se utilizan datos provinciales, pero la metodología es robusta y podría aplicarse a desagregaciones más detalladas, como comarcas o áreas metropolitanas, aprovechando las ventajas de disponer de datos de alta frecuencia.

Para lograrlo, el resto del trabajo se organiza de la siguiente manera. En la próxima sección se presentan los datos que se emplearán. En la sección tercera se expone la metodología de Phillips y Sul (2007), tanto para contrastar la hipótesis nula de convergencia (equivalente

a la existencia de un patrón de comportamiento común en los datos) como para analizar la posible existencia de clubes de convergencia. La sección cuarta presenta los resultados más destacados obtenidos al aplicar esta metodología. El trabajo termina con una revisión de las conclusiones más destacadas alcanzadas.

2. DATOS

Los datos utilizados han sido obtenidos del Instituto Nacional de Estadística, específicamente los índices de precios al consumo (IPC) mensuales para las 52 provincias españolas. La muestra abarca el período 2002:1-2023:9 y considera tanto el valor total del IPC como el de sus 12 grupos ECOICOP (*European Classification of Individual Consumption by Purpose*). De este modo, se descompuso el comportamiento total de los precios en componentes como (G01) Alimentos y bebidas no alcohólicas; (G02) Bebidas alcohólicas y tabaco; (G03) Vestido y calzado; (G04) Vivienda; (G05) Menaje; (G06) Medicina; (G07) Transporte; (G08) Comunicaciones; (G09) Ocio y cultura; (G10) Enseñanza; (G11) Hoteles, cafés y restaurantes y (G12) Otros bienes y servicios.

Dado que los datos están expresados en el año base 2021 y el objetivo es analizar posibles disparidades, se cambió la base a 2002, el inicio de la muestra, para evitar conclusiones espurias. Este ajuste aborda un problema detectado en el artículo de Phillips y Sul (2007), cuya metodología se sigue en este trabajo. Hay que tener en cuenta que, de otra manera, los índices de precios toman valores similares al final de la muestra, generando una falsa idea de convergencia. Después de este cambio de base, se presenta una breve descripción de los datos para una comprensión más clara. Los cuadros 1-3 muestran las tasas de crecimiento para varios períodos de interés, las 52 provincias españolas y la economía española en su totalidad.

Cuadro 1.

Tasas de crecimiento IPC total

Provincia	90203	90219	92023	92223
Álava	0,19	0,17	0,34	0,37
Albacete	0,19	0,16	0,38	0,47
Alicante	0,18	0,15	0,37	0,43
Almería	0,19	0,16	0,35	0,42
Ávila	0,19	0,16	0,40	0,48
Badajoz	0,17	0,14	0,38	0,43
Islas Baleares	0,19	0,16	0,37	0,44
Barcelona	0,21	0,19	0,32	0,38
Burgos	0,19	0,16	0,35	0,41
Cáceres	0,17	0,15	0,33	0,37

Cuadro 1. (continuación)

Tasas de crecimiento IPC total

<i>Provincia</i>	<i>90203</i>	<i>90219</i>	<i>92023</i>	<i>92223</i>
Cádiz	0,18	0,14	0,37	0,46
Castellón	0,19	0,15	0,39	0,45
Ciudad Real	0,19	0,16	0,39	0,45
Córdoba	0,19	0,16	0,38	0,41
La Coruña	0,19	0,17	0,37	0,41
Cuenca	0,19	0,15	0,38	0,45
Gerona	0,20	0,18	0,35	0,41
Granada	0,18	0,15	0,37	0,45
Guadalajara	0,19	0,15	0,39	0,45
Guipúzcoa	0,19	0,16	0,35	0,42
Huelva	0,19	0,15	0,38	0,45
Huesca	0,19	0,16	0,37	0,41
Jaén	0,18	0,15	0,37	0,43
León	0,20	0,16	0,41	0,47
Lérida	0,21	0,18	0,38	0,45
La Rioja	0,19	0,16	0,36	0,42
Lugo	0,19	0,16	0,37	0,43
Madrid	0,18	0,16	0,31	0,36
Málaga	0,20	0,16	0,38	0,46
Murcia	0,19	0,16	0,36	0,43
Navarra	0,19	0,16	0,37	0,44
Ourense	0,18	0,15	0,38	0,44
Asturias	0,18	0,15	0,34	0,41
Palencia	0,18	0,15	0,36	0,42
Las Palmas	0,16	0,12	0,33	0,42
Pontevedra	0,19	0,16	0,39	0,46
Salamanca	0,18	0,15	0,35	0,42
Santa Cruz de Tenerife	0,16	0,12	0,36	0,45
Cantabria	0,19	0,16	0,36	0,42
Segovia	0,19	0,16	0,38	0,42
Sevilla	0,18	0,15	0,36	0,43
Soria	0,19	0,16	0,38	0,43

Cuadro 1. (continuación)**Tasas de crecimiento IPC total**

<i>Provincia</i>	<i>90203</i>	<i>90219</i>	<i>92023</i>	<i>92223</i>
Tarragona	0,18	0,15	0,35	0,39
Teruel	0,19	0,16	0,36	0,44
Toledo	0,19	0,15	0,41	0,46
Valencia	0,18	0,16	0,34	0,40
Valladolid	0,19	0,16	0,35	0,40
Vizcaya	0,19	0,17	0,36	0,42
Zamora	0,19	0,16	0,40	0,47
Zaragoza	0,18	0,16	0,35	0,38
Ceuta	0,17	0,14	0,36	0,45
Melilla	0,19	0,15	0,43	0,48

g_{xy} es la tasa promedio de crecimiento entre los periodos 20xx y 20yy

Fuentes: Elaboración propia.

En el **cuadro 1**, que se centra en el comportamiento del IPC total, se observa cierta heterogeneidad entre las provincias en cuanto al crecimiento promedio. El crecimiento mensual promedio del IPC de la economía española fue del 0,19 %. Las provincias de Lérida y Barcelona destacan con un crecimiento del 0,21 %, mientras que Las Palmas y Santa Cruz de Tenerife tienen un crecimiento inferior. Si se considera el período anterior al COVID-19 (2002:1-2019:12), los resultados son ligeramente diferentes. En primer lugar se observa que el crecimiento promedio de los precios para el total de la economía española fue del 0,16 %, si bien para las provincias de Lérida, Gerona y Barcelona fue ligeramente superior, mientras que para las ya mencionadas provincias de Las Palmas y de Santa Cruz de Tenerife fue netamente inferior (0,12 %).

El comportamiento pos-COVID-19 es realmente diferente. Podemos observar cómo el crecimiento promedio del período 2020:1-2023:9 fue de un 0,35 % para el total de la economía española, casi el doble de lo observado en el período precedente. Destacan los crecimientos de Melilla (0,43 %), León y Toledo, ambos con un 0,41 %, así como los de Madrid (0,31 %) y Barcelona (0,32 %). Incluso, si consideramos el período 2022:1-2023:9 las diferencias son más notables, por cuanto el crecimiento promedio de los precios de la economía española fue de un 0,41 %. Las mayores alzas en los precios ocurrieron en las provincias de Melilla y Ávila (0,48 %), mientras que las provincias donde los precios presentaron menor crecimiento promedio fueron las de Madrid (0,36 %), Álava y Cáceres (ambas con un 0,37 %).

Estos resultados se pueden complementar con el análisis de la **figura 1** en el que se presenta la evolución temporal del coeficiente de variación del IPC total. En él se puede apreciar

cómo la dispersión muestral crece a lo largo de la muestra, si bien desde finales de 2019 hasta mediados de 2020 existe un claro retroceso en la dispersión. Posteriormente, vuelve a retomar su senda alcista hasta mediados de 2022, disminuyendo significativamente a partir de este período.

Figura 1.

σ -convergencia. Coeficiente de variación IPC total



Fuente: Elaboración propia.

Si en lugar de considerar el IPC total analizamos el comportamiento de los 12 grupos ECOICIP podemos encontrar nuevos resultados. En el **cuadro 2** se presentan los crecimientos promedios desde el año 2002 hasta el final de la muestra, respectivamente, para cada una de las provincias y de los grupos que hemos considerados.

Cuadro 2.

Tasa de crecimiento promedio de la muestra de los grupos ECOICOP y del IPC total

Provincia	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11	G12	Total
Álava	0,28	0,33	0,06	0,21	0,13	0,09	0,21	-0,08	0,03	0,19	0,22	0,20	0,19
Albacete	0,24	0,36	0,08	0,24	0,13	0,04	0,24	-0,07	0,00	0,27	0,25	0,19	0,19
Alicante	0,24	0,35	0,09	0,20	0,12	0,03	0,24	-0,10	0,01	0,21	0,23	0,19	0,18
Almería	0,26	0,36	0,07	0,21	0,08	0,08	0,23	-0,06	0,00	0,21	0,20	0,17	0,19
Ávila	0,25	0,34	0,02	0,26	0,11	0,09	0,23	-0,08	0,01	0,21	0,27	0,19	0,19
Badajoz	0,25	0,36	0,03	0,22	0,08	0,02	0,21	-0,10	-0,05	0,23	0,21	0,18	0,17
Islas Baleares	0,24	0,32	0,07	0,21	0,11	0,07	0,23	-0,07	0,03	0,21	0,25	0,20	0,19

Cuadro 2. (continuación)

Tasa de crecimiento promedio de la muestra de los grupos ECOICOP y del IPC total

<i>Provincia</i>	<i>G01</i>	<i>G02</i>	<i>G03</i>	<i>G04</i>	<i>G05</i>	<i>G06</i>	<i>G07</i>	<i>G08</i>	<i>G09</i>	<i>G10</i>	<i>G11</i>	<i>G12</i>	<i>Total</i>
Barcelona	0,26	0,34	0,11	0,22	0,17	0,10	0,23	-0,10	0,08	0,25	0,24	0,23	0,21
Burgos	0,26	0,33	0,06	0,21	0,16	0,10	0,22	-0,07	0,01	0,21	0,22	0,18	0,19
Cáceres	0,22	0,35	0,02	0,20	0,10	0,06	0,22	-0,09	-0,03	0,18	0,21	0,19	0,17
Cádiz	0,25	0,35	0,04	0,20	0,09	0,07	0,23	-0,12	-0,02	0,26	0,23	0,17	0,18
Castellón	0,26	0,34	0,11	0,22	0,09	0,01	0,24	-0,11	-0,02	0,22	0,22	0,19	0,19
Ciudad Real	0,25	0,36	0,10	0,24	0,07	0,04	0,24	-0,08	-0,07	0,19	0,23	0,18	0,19
Córdoba	0,24	0,33	0,08	0,21	0,10	0,10	0,22	-0,09	0,00	0,23	0,24	0,18	0,19
La Coruña	0,25	0,31	0,09	0,24	0,15	0,10	0,23	-0,09	0,01	0,19	0,25	0,17	0,19
Cuenca	0,24	0,36	0,06	0,27	0,12	0,06	0,22	-0,08	-0,04	0,21	0,23	0,18	0,19
Gerona	0,26	0,32	0,11	0,24	0,13	0,11	0,24	-0,11	0,04	0,25	0,24	0,23	0,20
Granada	0,26	0,35	0,06	0,23	0,12	0,03	0,22	-0,07	-0,02	0,24	0,24	0,18	0,18
Guadalajara	0,26	0,36	0,04	0,22	0,11	0,10	0,22	-0,09	0,05	0,23	0,21	0,16	0,19
Guipúzcoa	0,25	0,31	0,03	0,22	0,12	0,10	0,24	-0,08	0,06	0,20	0,24	0,19	0,19
Huelva	0,24	0,35	0,06	0,22	0,12	0,09	0,23	-0,08	-0,01	0,23	0,21	0,16	0,19
Huesca	0,25	0,31	0,07	0,25	0,13	0,10	0,23	-0,06	-0,02	0,22	0,22	0,19	0,19
Jaén	0,25	0,34	0,07	0,22	0,08	0,08	0,24	-0,07	-0,04	0,23	0,21	0,15	0,18
León	0,26	0,33	0,11	0,25	0,15	0,10	0,22	-0,09	-0,03	0,21	0,26	0,19	0,20
Lérida	0,26	0,33	0,11	0,24	0,15	0,11	0,23	-0,07	0,04	0,24	0,24	0,22	0,21
La Rioja	0,24	0,33	0,09	0,23	0,16	0,09	0,23	-0,10	0,05	0,21	0,22	0,21	0,19
Lugo	0,24	0,30	0,07	0,21	0,14	0,08	0,23	-0,12	0,07	0,22	0,26	0,16	0,19
Madrid	0,25	0,35	0,08	0,19	0,12	0,04	0,22	-0,08	0,05	0,21	0,22	0,21	0,18
Málaga	0,25	0,36	0,05	0,24	0,13	0,06	0,23	-0,07	-0,02	0,24	0,26	0,19	0,20
Murcia	0,26	0,35	0,09	0,21	0,11	0,04	0,23	-0,09	0,02	0,22	0,23	0,19	0,19
Navarra	0,24	0,33	0,10	0,22	0,13	0,14	0,22	-0,10	0,05	0,22	0,22	0,20	0,19
Ourense	0,24	0,31	0,05	0,24	0,12	0,09	0,22	-0,10	0,02	0,23	0,23	0,15	0,18
Asturias	0,23	0,33	0,10	0,23	0,11	0,04	0,22	-0,10	0,00	0,20	0,23	0,19	0,18
Palencia	0,24	0,35	0,02	0,24	0,11	0,07	0,22	-0,10	-0,01	0,22	0,21	0,18	0,18
Las Palmas	0,22	0,36	-0,01	0,16	0,04	0,05	0,23	-0,11	-0,03	0,25	0,20	0,13	0,16
Pontevedra	0,24	0,33	0,07	0,24	0,15	0,10	0,23	-0,09	0,01	0,19	0,27	0,18	0,19
Salamanca	0,26	0,34	0,03	0,21	0,08	0,05	0,23	-0,10	-0,03	0,20	0,22	0,17	0,18
Santa Cruz de Tenerife	0,25	0,35	-0,06	0,16	0,04	0,02	0,27	-0,06	-0,01	0,18	0,19	0,15	0,16

Cuadro 2. (continuación)

Tasa de crecimiento promedio de la muestra de los grupos ECOICOP y del IPC total

Provincia	G01	G02	G03	G04	G05	G06	G07	G08	G09	G10	G11	G12	Total
Cantabria	0,25	0,34	0,06	0,23	0,10	0,08	0,23	-0,06	0,03	0,17	0,23	0,20	0,19
Segovia	0,26	0,33	0,11	0,23	0,08	0,10	0,22	-0,10	0,06	0,26	0,23	0,19	0,19
Sevilla	0,25	0,35	0,07	0,21	0,12	0,04	0,21	-0,11	-0,01	0,23	0,22	0,19	0,18
Soria	0,25	0,33	0,09	0,25	0,10	0,09	0,23	-0,06	0,00	0,20	0,23	0,16	0,19
Tarragona	0,24	0,34	0,03	0,21	0,07	0,09	0,24	-0,08	-0,01	0,19	0,22	0,20	0,18
Teruel	0,25	0,34	0,07	0,27	0,11	0,13	0,24	-0,09	-0,01	0,17	0,23	0,18	0,19
Toledo	0,25	0,35	0,07	0,26	0,08	0,09	0,21	-0,06	-0,01	0,24	0,23	0,18	0,19
Valencia	0,25	0,35	0,07	0,24	0,11	0,06	0,22	-0,12	0,03	0,18	0,25	0,18	0,18
Valladolid	0,24	0,35	0,08	0,22	0,14	0,09	0,22	-0,07	0,01	0,22	0,22	0,17	0,19
Vizcaya	0,26	0,33	0,11	0,23	0,17	0,09	0,22	-0,09	0,03	0,23	0,23	0,20	0,19
Zamora	0,27	0,35	0,08	0,23	0,11	0,09	0,23	-0,06	0,00	0,18	0,22	0,16	0,19
Zaragoza	0,24	0,34	0,07	0,23	0,12	0,10	0,22	-0,08	0,02	0,24	0,23	0,20	0,18
Ceuta	0,23	0,28	0,08	0,19	0,06	0,02	0,28	-0,05	0,01	0,14	0,19	0,15	0,17
Melilla	0,28	0,27	0,04	0,21	0,10	0,04	0,30	-0,09	-0,01	0,21	0,22	0,15	0,19
España	0,25	0,34	0,08	0,22	0,12	0,07	0,23	-0,09	0,03	0,22	0,23	0,20	0,19

Fuentes: Elaboración propia.

Si comenzamos analizando el **cuadro 2**, vemos que las tasas de crecimiento promedio son bastante heterogéneas. Si consideramos la media nacional, se observa que el menor crecimiento durante el período fue el del grupo 08 (Comunicaciones), con un crecimiento negativo promedio del -0,09 %. Frente a ello, el grupo 02 (Bebidas alcohólicas y tabaco) tiene un crecimiento promedio del 0,34 %. También hay que señalar que el grupo 01 (Alimentos y bebidas no alcohólicas), presenta el segundo crecimiento promedio más elevado (0,25 %).

Si desagregamos por provincias, de nuevo se observa gran disparidad en los datos. En particular, el coeficiente de variación de las tasas de crecimiento promedio varía entre el 4,9 % del grupo 01 (Alimentos y bebidas no alcohólicas) y el 49,5 % del grupo 04 (Vestido y calzado). El máximo crecimiento promedio se encuentra en el grupo 02 (Bebidas alcohólicas y tabaco) en la provincia de Albacete, mientras que la menor tasa de crecimiento promedio es la del grupo 08 (Comunicaciones) en la provincia de Lugo (-0,12 %).

En la **figura 2** se presenta la evolución de la dispersión para los diversos componentes del IPC. Como se puede observar, el comportamiento es claramente creciente para todos los grupos considerados, si bien el crecimiento promedio no es homogéneo. Así, los grupos 07 (Transporte) y 01 (Alimentos y bebidas no alcohólicas), presentan los menores crecimientos

promedios desde el año 2006 (1,4 % y 1,8 %), respectivamente. Por el contrario, los grupos 09 (Ocio y cultura), 05 (Muebles, artículos del hogar y artículos para el mantenimiento corriente del hogar) y 12 (Otros bienes y servicios) tienen crecimientos promedios por encima del 3 % (3,2 %, 3,2 % y 3,5 %, respectivamente). Además, se puede comprobar que los máximos valores de la dispersión se producen mayoritariamente alrededor del final de la muestra. Son claras excepciones los casos de los grupos 02 (Bebidas alcohólicas y tabaco), 04 (Vivienda, agua, electricidad, gas y otros combustibles) y 07 (Transporte), este último en menor medida. En todo caso, el hecho de presentar tendencias crecientes en la evolución del coeficiente de dispersión supone que existe cierta evidencia en contra de la existencia de σ -convergencia.

Figura 2.

σ -convergencia. Coeficiente de variación de los grupos del IPC



Fuente: Elaboración propia.

A partir de esta información inicial es sencillo concluir que la evolución de los precios en España no ha sido homogénea ni desde un punto de vista geográfico, ni desde una perspectiva temporal. Por lo que no resulta sencillo pensar que la evolución provincial de los precios en España exhiba un único patrón de comportamiento. Por el contrario, es más que probable que podemos concluir que existen disparidades significativas.

No obstante, los métodos utilizados en esta sección son meramente descriptivos, por lo que parece adecuado utilizar herramientas que sean capaces de ofrecer conclusiones más

robustas y, sobre todo, que permitan el uso de un número importante de datos, como los que aquí se usan. La metodología a emplear se presenta en la siguiente sección.

3. METODOLOGÍA

Antes de presentar la metodología que vamos a emplear, creemos necesario recordar la relación entre los análisis de convergencia y con los conceptos de β -convergencia y σ -convergencia, inicialmente propuestos por Barro y Sala-i-Martin (1990) al estudiar la dispersión del PIB per cápita. La β -convergencia sugiere que las economías menos prósperas tienden a crecer más rápidamente que las más ricas, reduciendo la brecha entre ellas, mientras que la σ -convergencia implica una disminución en la dispersión transversal de la variable a lo largo del tiempo.

Ambos métodos han recibido severas críticas; por ejemplo, la β -convergencia ha sido señalada por indicar convergencia incluso en casos donde podría no existir genuinamente, según De Long (1988) y Quah (1993). Además, ninguno de estos métodos prueba explícitamente la hipótesis nula de convergencia, lo que ha llevado a la propuesta de enfoques alternativos en la literatura.

Para solucionar este problema, Phillips y Sul (2007, 2009) desarrollaron una metodología asociada al concepto de σ -convergencia, que es la que se sigue en este trabajo. Estos autores diseñan un estadístico para contrastar la hipótesis nula de convergencia y determinar la existencia de clubes de convergencia si esta hipótesis es rechazada.

La metodología de Phillips-Sul resulta muy adecuada, en especial en un entorno de uso masivo de datos, por cuanto es muy flexible con respecto a las características temporales de las variables. Además, permite aprovechar las ventajas que ofrece el uso combinado de datos de series temporales para un número amplio de corte transversales. En consecuencia, supera claramente al análisis clásico de convergencia. Por lo tanto, no es sorprendente que la metodología esté ganando popularidad y se aplique ampliamente para analizar disparidades en la evolución de diversas variables. En concreto, en el caso de los precios hay que reseñar que el propio artículo original de Phillips y Sul (2007) estudia la evolución de los precios en los Estados Unidos. Otros ejemplos son los trabajos de Liu y Yeh (2012) quienes estudian los precios en las ciudades USA, Lin y Robberts (2023) que estudian el caso de los precios de la vivienda en el Reino Unido o Gil *et al.* (2023) quienes estudian los precios del mercado internacional de la madera aserrada, entre otros casos.

Siguiendo a Phillips-Sul, denominemos nuestra variable de interés (el logaritmo del índice de precios al consumo total y desglosado en sus doce grupos) como X_{it} , donde i representa las 52 provincias españolas, y t abarca el período 2006:1-2023:9. Como se puede apreciar, la muestra utilizada no se corresponde con el total de la muestra disponible. Phillips-Sul aconsejan eliminar algunas observaciones iniciales en el caso de los precios, para evitar el efecto base, ya que todas las provincias toman el valor 100 al comienzo de la muestra.

Entonces, podemos desagregar la variable de interés de la siguiente manera $X_{it} = \delta_{it}\mu_t$ donde δ_{it} es el componente idiosincrático y μ_t es el componente de tendencia común.

Dado que el número de incógnitas en el modelo es mayor que el número de observaciones en el panel, Phillips y Sul (2007) eliminan el factor común, obteniendo la variable h_{it} .

$$h_{it} = \frac{X_{it}}{N^{-1} \sum_{i=1}^N X_{it}} = \frac{\delta_{it}\mu_t}{N^{-1} \sum_{i=1}^N \delta_{it}\mu_t} = \frac{\delta_{it}}{N^{-1} \sum_{i=1}^N \delta_{it}} \quad [1]$$

Esta variable h_{it} representa el parámetro relativo de transición (transition path) y mide la posición relativa de cada corte transversal con el promedio del panel en el período t . Por definición, la media transversal de h_{it} es la unidad. Por tanto, cuando existe convergencia, h_{it} debería aproximarse a la unidad para todos los i cuando t tiende a infinito.

La varianza transversal de h_{it} se define de la siguiente manera:

$$H_t = \frac{1}{N} \sum_{i=1}^N (h_{it} - 1)^2 \quad [2]$$

Esta varianza, bajo el supuesto de convergencia, tiende hacia 0 conforme la muestra crece. Por tanto, para analizar la presencia de convergencia basta con estudiar si esta varianza decrece hacia 0. Para ello, Phillips y Sul (2007) proponen estimar el siguiente modelo:

$$\ln \frac{H_1}{H_t} - \ln L(t) = a + \beta \ln L(t) + u_t, \quad t = [rT], [rT] + 1, \dots, T \quad \text{con } r > 0 \quad [3]$$

donde la variable dependiente del modelo está claramente relacionada con la varianza de sección-cruzada, de manera que se puede estudiar si la varianza del panel (H_t) tiene tendencia decreciente a partir del valor del parámetro β . Para mejorar las propiedades del método propuesto introducen $L(t)$ que es una función de t . A través de diversos ejercicios de simulación, los autores eligen usar $L(t) = \ln t$. De igual manera, r es el primer porcentaje de los datos que se descarta. Según los experimentos de Monte Carlo realizados por Phillips y Sul (2007), r debería tomar valores en el intervalo (0,2, 0,3) dependiendo del tamaño de la muestra. En nuestro caso se ha usado el valor 1/3.

De acuerdo a este modelo, existe convergencia siempre que $\beta \geq 0$, mientras que si $\beta < 0$, entonces tenemos un comportamiento divergente. Para distinguir entre estas dos alternativas, PS proponen usar el denominado log t-ratio que no es otra cosa que el t-ratio para contrastar la hipótesis nula $\beta \geq 0$ en el modelo anterior. Estos autores recomiendan emplear métodos de estimación robustos a la presencia de autocorrelación y heterocedasticidad. Entonces, si el log t-ratio es inferior a -1,65, se rechaza la hipótesis nula de convergencia. No obstante, hay que tener en cuenta que podrían existir los denominados clubes de convergencia. Para tener en cuenta esta posibilidad, PS desarrollan un algoritmo para estimar dichos clubes de convergencia. El algoritmo consta de los siguientes pasos:

- Ordenar los miembros del panel según su último valor de la variable de interés. En este caso, se ordenan de mayor a menor las provincias en función del índice de precios del último período de la muestra.
- Efectuada esta ordenación, se debe generar un grupo de provincias inicial. Para ello, se van agregando provincias en función de la ordenación anterior mientras que el $\log t$ -ratio sea inferior a $-1,65$.
- Una vez formado este grupo inicial, se prueba con el resto de las provincias no incluidas y se reestima el modelo [3] para cada provincia añadida. Si el $\log t$ -ratio es inferior a $-1,65$, dicha provincia se descarta. En el caso contrario, la provincia se añade al grupo inicial. Una vez consideradas todas las provincias, se tiene formado el primer club de convergencia.
- Ahora, debemos analizar la posible existencia de múltiples clubes de convergencia. Por lo que se repiten los pasos anteriores para las provincias descartadas. Si alguna provincia no puede incluirse en ningún club, se considera que diverge y no se incluye en ningún club de convergencia.

Phillips y Sul (2007) señalan la posibilidad de encontrar más clubes de convergencia de los que realmente existen debido a la naturaleza conservadora del algoritmo. Para eliminar la sobreestimación de clubes convergentes, recomiendan realizar pruebas de regresión $\log t$ entre clubes adyacentes para determinar si deben fusionarse en clubes más grandes. Este procedimiento de fusión se aplica hasta que no se pueden fusionar más clubes.

Finalmente, Phillips y Sul (2007) aconsejan aplicar el filtro de Hodrick y Prescott (1997) para eliminar cierto ruido de la variable y mejorar la potencia del estadístico. En nuestro caso, no se ha utilizado este filtro dado que la evidencia en contra de la hipótesis nula era notable y, además, debería haberse usado un valor elevado del parámetro de alisado dado que nuestros datos son mensuales. Este filtro podría distorsionar los resultados y, por tanto, hemos preferido usar los datos originales.

4. RESULTADOS

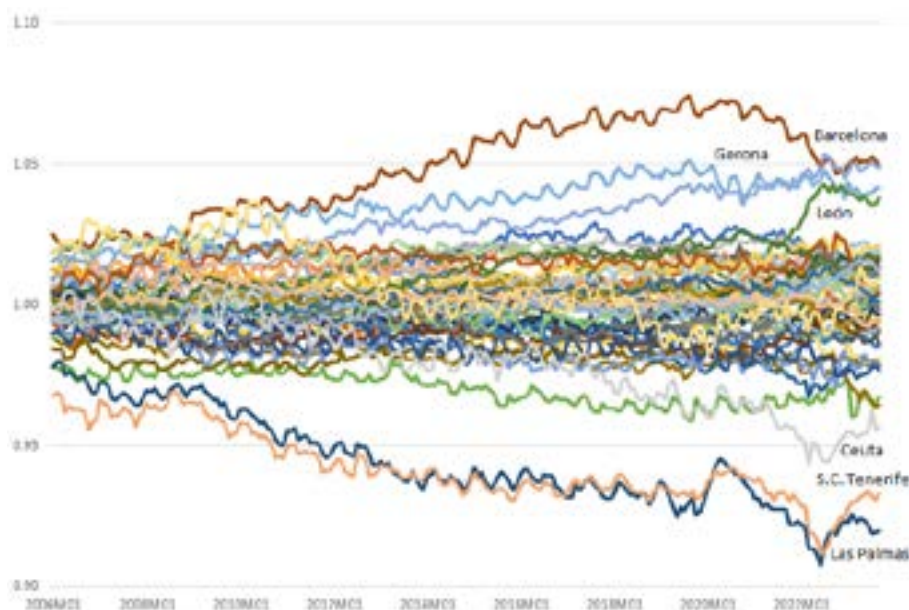
En primer lugar, hay que tener en cuenta el estadístico $\log t$ para contrastar la hipótesis nula de convergencia. El valor de este estadístico es $-28,81$, inferior al valor crítico ($-1,65$), por lo que la evidencia en contra de la hipótesis nula es clara. En consecuencia, se puede afirmar que la evolución de los precios provinciales en España presenta diferencias significativas y no existe un único patrón de comportamiento.

Este hecho se puede confirmar analizando la [figura 3](#), donde se presentan los valores de la variable h , definida en [1] y que habitualmente es denominada *transition path*, para las diversas provincias. Esta variable proporciona una medida de cuál es la distancia de cada una de las variables con respecto al promedio del panel. En la mencionada figura se observa

que existen provincias cuyos valores h del final de la muestra se agrupan en torno a valores claramente alejados de la unidad. Es el caso de Barcelona, Lérida, Gerona y León, con valores superiores a 1,038. Por el contrario, las provincias de Las Palmas y de Santa Cruz de Tenerife muestran valores ligeramente superiores a 0,92, mientras que el de Ceuta se sitúa en 0,96. Por lo tanto, parece que existen diversos patrones de comportamiento y que las provincias se pueden agrupar en torno a ellos. Para ello, parece conveniente utilizar el algoritmo propuesto por Phillips y Sul (2007) y analizar si pueden existir distintos clubes de convergencia.

Figura 3.

Transition paths



Fuente: Elaboración propia.

Los resultados de aplicar el citado algoritmo nos llevan a estimar la existencia de cinco clubes de convergencia. La composición provincial de cada uno de estos clubes de convergencia es la siguiente:

- Club 1: Barcelona, Gerona, León, Lérida y Málaga.
- Club 2: Álava, Albacete, Ávila, Islas Baleares, Burgos, Castellón, Ciudad Real, Córdoba, La Coruña, Cuenca, Guadalajara, Guipúzcoa, Huelva, Huesca, La Rioja, Lugo, Murcia, Navarra, Pontevedra, Cantabria, Segovia, Soria, Teruel, Toledo, Valladolid, Vizcaya, Zamora y Melilla.
- Club 3: Alicante, Almería, Granada, Orense, Asturias, Sevilla, Valencia y Zaragoza.

- Club 4: Badajoz, Cáceres, Cádiz, Jaén, Madrid, Palencia, Salamanca y Tarragona.
- Club 5: Las Palmas, Santa Cruz de Tenerife y Ceuta.

Si comparamos los resultados con el análisis de la **figura 3**, resulta directo asociar los clubes 1 y 5 con la evolución de los valores de la variable h . Para entender mejor la composición de los clubes, la **figura 4** presenta la composición de los diferentes clubes de convergencia estimados en forma de mapa. La pertenencia de cada una de las provincias a los diferentes clubes estimados se caracteriza de un color, siendo los colores más cálidos (rojo y naranja) los primeros clubes, asociados a los valores más altos de los precios al final de la muestra, mientras que las provincias incluidas en los últimos clubes se colorean en tonos verdosos.

Figura 4.

Clubes de convergencia estimados



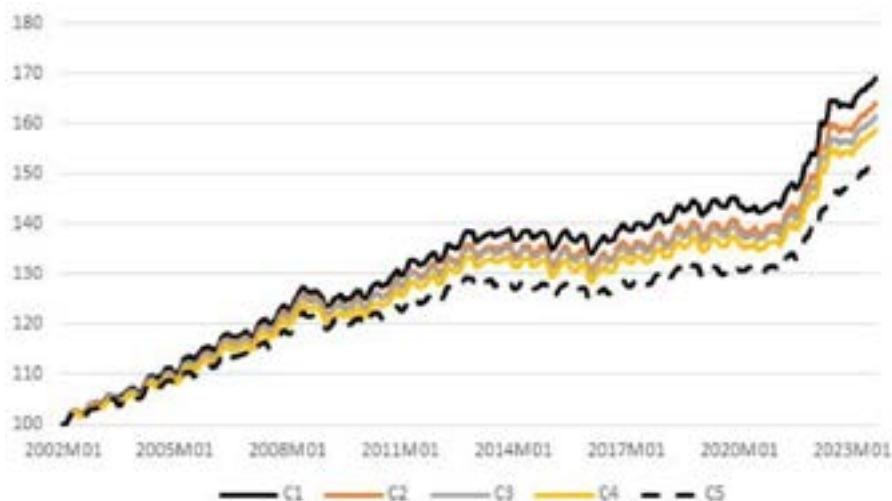
Fuente: Elaboración propia.

Como se puede apreciar en dicha figura, existe cierto grado de correlación espacial. Por ejemplo, las provincias incluidas en el club 1 están situadas en Cataluña, si bien hay que añadir a éstas los casos de León y Málaga. Frente a ellas, tenemos el caso de las provincias insulares canarias, que pertenece al club 5, conjuntamente con Ceuta. También es interesante resaltar que el club 2 es el que contiene el mayor número de provincias, mientras que cuatro de las cinco capitales de provincias más pobladas (Madrid, Sevilla, Valencia y Zaragoza) se encuentran en los clubes 3 y 4. Por el contrario, provincias con capitales escasamente pobladas, como Huesca y Teruel, están en el club 2, al igual que las provincias que rodean a Madrid.

La **figura 5** presenta la media aritmética del índice de precios para las provincias incluidas en cada club estimado. Como se puede apreciar, las provincias del club 1 muestran las medias más altas, mientras que las incluidas en el club 5 son las más bajas. En todo caso, los perfiles de la evolución del IPC son muy similares para cada uno de los clubes estimados, mostrando un claro intervalo de escaso crecimiento durante el período 2012-2021. A partir de dicho año, los precios experimentan un claro período alcista, cuya fuerza llega hasta el final de la muestra utilizada.

Figura 5.

Evolución de los comportamientos medios de los clubes de emergencia. IPC total



Fuente: Elaboración propia.

El análisis descriptivo realizado con anterioridad mostró que la evolución de los diferentes componentes del índice de precios no era homogénea. Por ello, cabe plantearse si los clubes estimados son también válidos para cada uno de los 12 grupos en los que se puede desagregar el índice de precios al consumo. En el **cuadro 3** se presentan los valores del log t-ratio para contrastar la hipótesis nula de convergencia en cada uno de estos grupos. Para ello, hemos estimado la ecuación [3] para cada uno de estos grupos y hemos obtenido el correspondiente log t-ratio para contrastar la hipótesis nula de convergencia. Se observa que dicha hipótesis se rechaza mayoritariamente, por lo que existe evidencia de disparidad en la evolución de los precios para las diferentes provincias. No obstante, también se ve que esta hipótesis nula no puede rechazarse para los grupos G02 (Bebidas alcohólicas y tabaco) y G07 (Transporte), por lo que para estos dos grupos, la evolución de los precios provinciales ha sido estadísticamente idéntica.

Para el resto de los grupos, tal y como hemos comentado, se rechaza la hipótesis nula de convergencia y, por tanto, no existe un único comportamiento. Pero, al igual que para el total

Cuadro 3.

Contraste de convergencia para los grupos provinciales del IPC

<i>Grupos</i>	<i>log t-ratio</i>
G01. Alimentos y bebidas no alcohólicas	-0,94 (-16,84) *
G02. Bebidas alcohólicas y tabaco	-0,03 (-0,29)
G03. Vestido y calzado	-1,01 (-6,87)
G04. Vivienda, agua, electricidad, gas y otros combustibles	-0,45 (-3,05)
G05. Muebles, artículos del hogar y artículos para el mantenimiento corriente del hogar	-1,43 (-46,53)
G06. Sanidad	-0,90 (-12,01) *
G07. Transporte	-0,17 (-1,31)
G08. Comunicaciones	-1,36 (-14,97)
G09. Ocio y cultura	-1,20 (-22,57)
G10. Enseñanza	-0,54 (-5,07)
G11. Restaurantes y hoteles	-0,91 (-7,79)
G12. Otros bienes y servicios	-1,34 (-39,51)

Notas: La segunda columna presenta el log t-ratio para contrastar la hipótesis de convergencia entre paréntesis. *significa que no se puede rechazar la hipótesis nula de convergencia.

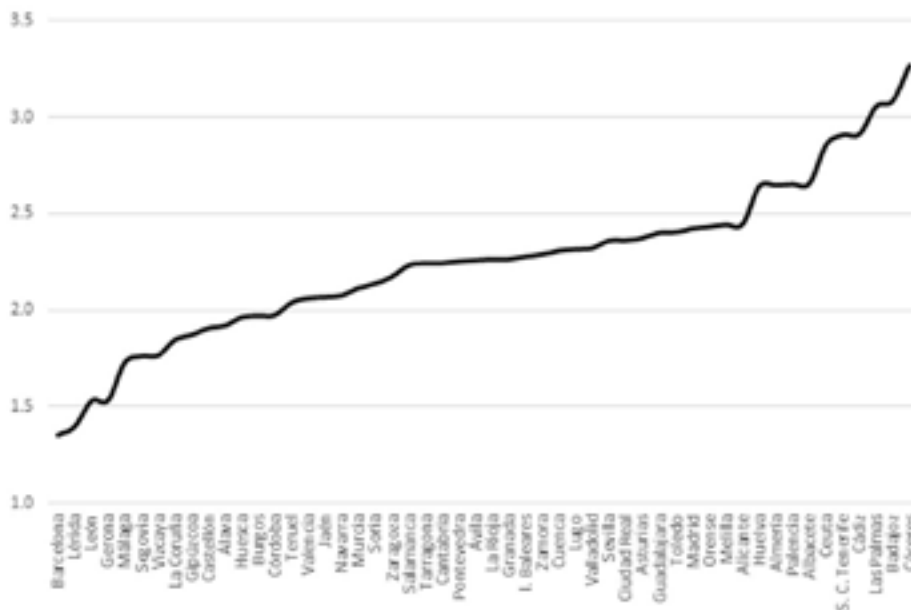
Fuentes: Elaboración propia.

del índice de precios al consumo, parece adecuado analizar la posible presencia de clubes de convergencia. Para ello, hemos aplicado el algoritmo de Phillips-Sul, encontrando diversos resultados. Todos ellos quedan resumidos en la **figura 6** en la que se presenta el valor promedio del ordinal del club de convergencia en el que cada provincia se ha incluido, multiplicado por la ponderación de cada uno de los grupos en el índice de precios de dicha provincia. Como se puede apreciar, los valores de las provincias de Barcelona, Lérida, Gerona y León, todas ellas incluidas en el club 1 para el total del índice de precios al consumo, son los más bajos. Por tanto, dichas provincias muestran valores del índice de precios al consumo elevados para cada uno de los componentes, no solamente para algunos de sus componentes. Del mismo modo, las provincias de Ceuta, Santa Cruz de Tenerife, Cádiz, Las Palmas, Badajoz y

Cáceres son las que presentan un valor medio ponderado más elevado. Hay que señalar que dichas provincias se incluyeron en los clubes de convergencia 4 y 5 para el total del índice de precios al consumo.

Figura 6.

Club de convergencia de las provincias (ponderado por peso grupos del IPC)



Fuente: Elaboración propia.

Como caso particular, se analiza de forma separada la evolución de los precios del grupo 01 (Alimentación y Bebidas). Este grupo presenta el peso más elevado dentro del total del índice de precios al consumo. Esta ponderación va desde el valor 0,17 (Madrid) hasta el valor más alto, 0,25, para Ceuta.

El algoritmo de clusterización propuesto por Phillips-Sul estima la existencia de cuatro clubs de convergencia diferentes, cuya composición es la siguiente:

- Club 1: Álava, Almería, Barcelona, Castellón, Gerona, Guadalajara, León, Lérida, Salamanca, Segovia, Zamora y Melilla.
- Club 2: Albacete, Ávila, I. Baleares, Burgos, Ciudad Real, Córdoba, La Coruña, Cuenca, Granada, Guipúzcoa, Huesca, Jaén, Madrid, Málaga, Murcia, Palencia, Cantabria, Sevilla, Soria, Tarragona, Teruel, Toledo, Valencia, Valladolid, Vizcaya y Zaragoza.

- Club 3: Alicante, Badajoz, Cádiz, Huelva, La Rioja, Lugo, Navarra, Orense, Asturias, Pontevedra y Santa Cruz de Tenerife.
- Club 4: Cáceres, Las Palmas y Ceuta.

La **figura 7** presenta la composición geográfica de cada uno de los clubes. Hay que destacar que en el club 1 se mantienen las provincias de Barcelona, Gerona, Lérida y León, a las que se añaden Álava, Almería, Castellón, Guadalajara, Salamanca, Segovia, Zamora y Melilla. Por el contrario, el club 4 contiene las provincias con los precios más bajos: Cáceres, Las Palmas y Ceuta. Hay que resaltar que el club 2 agrupa el mayor número de provincias, incluyendo las que tienen capitales de provincias más pobladas, a excepción de Barcelona.

Figura 7.

Clubes de convergencia estimados. Grupo 01



Fuente: Elaboración propia.

La evolución promedio de los índices de precios del Grupo 01 se presenta en la **figura 8**. Se puede observar que el perfil es similar para los cuatro clubes de convergencia, aunque la tasa de crecimiento promedio del período es del 0,6 %, 0,25 %, 0,24 % y 0,22 % para, respectivamente, los clubes de convergencia 1, 2, 3 y 4. Si nos centramos en la parte final de la muestra, se observa un crecimiento netamente superior al visto con anterioridad. En concreto, las tasas de crecimiento promedio desde 2022:1 son del 0,99 %, 1,00 %, 1,04 % y 1,04 % para los clubes de convergencia 1, 2, 3 y 4, respectivamente. Por tanto, el período inflacionista ocurrido en este grupo de productos al final de la muestra ha sido muy similar para todas las provincias.

Figura 8.

Evolución promedio de los clubes de convergencia estimados. Grupo 01

Fuente: Elaboración propia.

5. CONCLUSIONES

La presente investigación ha analizado la evolución de los precios provinciales en España mediante la aplicación de técnicas que posibilitan manejar volúmenes significativos de información. Específicamente, la metodología propuesta por Phillips y Sul (2007) es la que hemos usado para estudiar la presencia de patrones de comportamiento compartidos en los índices de precios al consumo de las diversas provincias españolas durante el período 2002:1-2023:9.

Los resultados obtenidos revelan que la trayectoria de los precios a nivel provincial en España no sigue una pauta uniforme. Por el contrario, se observan disparidades significativas en su evolución a lo largo del tiempo para las distintas provincias. En particular, la aplicación de la mencionada metodología ha conducido al rechazo de la hipótesis nula de convergencia, evidenciando la existencia de cinco clubes de convergencia para el índice de precios al consumo provincial. La composición de estos clubes indica que provincias como Barcelona, Gerona, Lérida, León y Málaga exhiben precios más elevados, mientras que el grupo conformado por las provincias de Las Palmas, Santa Cruz de Tenerife y Ceuta presenta precios más bajos.

Estas disparidades persisten al desagregar el índice total en sus 12 componentes ECOICOP, donde se rechaza la hipótesis nula de convergencia en la mayoría de estos grupos, con excepción de los Grupos 02 (Bebidas alcohólicas y tabaco) y 07 (Transporte).

En resumen, este estudio resalta la complejidad y diversidad inherente a la evolución de los precios provinciales en España. La implementación de técnicas estadísticas que facilitan el procesamiento masivo de datos ha permitido identificar patrones de comportamiento comunes, proporcionando una comprensión más profunda de las dinámicas subyacentes en los precios españoles a nivel provincial. Estas conclusiones poseen relevancia en el ámbito de la toma de decisiones económicas, en especial si consideramos la formulación de políticas diferenciadas espacialmente.

Referencias

- BARRO, R. J. y SALA-I-MARTIN, X. (1990). Economic growth and convergence across the United States. *NBER Working Paper*, (w3419).
- BARRO, R. J. y SALA-I-MARTIN, X. (1992). Convergence. *Journal of political Economy*, 100(2), pp. 223-251.
- BAUMOL, W. J. (1986). Productivity growth, convergence, and welfare: what the long-run data show. *The American Economic Review*, 76(5), pp. 1072-1085.
- CHATTERJEE, A., GUPTA, S. y LAHIRI, S. N. (2015). On the residual empirical process based on the ALASSO in high dimensions and its functional oracle property. *Journal of Econometrics*, 186(2), pp. 317-324.
- GIL, J. M., MONTAÑÉS, A. y VÁSQUEZ-GONZÁLEZ, B. (2023). Are prices converging in the global sawnwood market? *Forest Policy and Economics*, 102998.
- HODRICK, R. J. y PRESCOTT, E. C. (1997). Postwar US business cycles: an empirical investigation. *Journal of Money, Credit, and Banking*, 29(1), pp. 1-16.
- DE LONG, J. B. (1988). Productivity growth, convergence, and welfare: comment. *The American Economic Review*, 78(5), pp. 1138-1154.
- LIN, P. T. y ROBERTS, A. (2023). Regional house price convergence: implications of monetary policy. *Regional Studies*, 58(5), pp. 1-13.
- LIU, W. H. y YEH, C. C. (2012). Convergence in price levels across US cities. *Economics Letters*, 114(3), pp. 245-248.
- PHILLIPS, P. C. y SUL, D. (2007). Transition modeling and econometric convergence tests. *Econometrica*, 75(6), pp. 1771-1855.
- PHILLIPS, P. C. y SUL, D. (2009). Economic transition and growth. *Journal of Applied Econometrics*, 24(7), pp. 1153-1185.
- QUAH, D. (1993). Galton's fallacy and tests of the convergence hypothesis. *The Scandinavian Journal of Economics*, 95(4), pp. 427-443.
- STOCK, J. H. y WATSON, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460), pp. 1167-1179.
- ZOU, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476), pp.1418-1429.