



COMUNICACIÓN

Título: Empresas granulares y desagregación regional: un análisis del caso español

Autores y e-mails de todos: Omar Blanco-Arroyo (blancoo@uji.es), David Vidal-Tómas (dvidal@uji.es), Alba Ruiz-Buforn (buforn@uji.es) y Simone Alfarano (alfarano@uji.es)

Departamento: Departamento de Economía

Universidad: Universitat Jaume I

Área Temática: *(indicar el área temática en la que se inscribe el contenido de la comunicación)*

Empresas, emprendimiento y dinámica empresarial en el territorio

Resumen: *(máximo 300 palabras)*

Siguiendo el planteamiento propuesto por Gabaix (2011), el presente trabajo tiene como objetivo evaluar la existencia de granularidad en las fluctuaciones del ciclo económico de las regiones españolas con mayor contribución al PIB español y, a su vez, con mayor concentración de grandes empresas. Estas regiones son: Comunidad de Madrid, Cataluña, País Vasco y Comunidad Valenciana. Las empresas granulares son aquellas que representan una proporción marginal del número total de una economía, pero, sin embargo, tienen un impacto significativo en las fluctuaciones en la tasa de crecimiento del PIB. Encontramos que País Vasco y Comunidad Valenciana son economías granulares. Comunidad de Madrid y Cataluña, sin embargo, no presentan un comportamiento granular. Por consiguiente, nuestro trabajo aporta evidencia que indica que puede existir comportamiento granular a nivel regional.

Palabras Clave: *(máximo 6 palabras)*

Granularidad, empresas granulares, perturbaciones idiosincrásicas, fluctuaciones agregadas, regiones.

Clasificación JEL: E32, C20



1. Introducción

El concepto de “scaling law”, un fundamento básico de los estudios de fenómenos críticos en física estadística (Stanley, 1971), fue observado por primera vez por Pareto (1896) en su trabajo empírico sobre la distribución estadística de la riqueza de los individuos más ricos en varios países. Pareto determinó empíricamente que la probabilidad de observar un individuo i con riqueza W_i mayor que un número w sigue una ley de potencia:

$$P(W_i > w) \propto 1/w^\alpha$$

donde α es una constante (véase el apartado 3). En lugar de ser una curiosidad estadística, este comportamiento de los datos empíricos se ha observado también en otros entornos económicos, como los mercados financieros, la población de las ciudades y el tamaño de las empresas (Gabaix, 2009).

En los años 90 y principio de la década de los 2000, en el campo de investigación conocido como “econofísica” (Mantegna y Stanley, 1999), varios físicos han descrito la aparición de estas regularidades empíricas como el “inprint” de fenómenos críticos. La economía como un sistema de unidades heterogéneas que interactúan ha sido el marco conceptual para poder aplicar estos conceptos en dicho ámbito interdisciplinar. La amplia disponibilidad de datos en los mercados financieros ha propiciado que sea el primer ámbito de aplicación de la econofísica; en particular, la caracterización de las fluctuaciones de los valores financieros. Numerosos modelos conectan las leyes de potencia empíricamente identificadas en las grandes variaciones de los precios de los activos con las estrategias adoptadas por agentes heterogéneos (Lux y Alfarano, 2016; Lux y Marchesi, 1999).

En macroeconomía, el interés por la ley de potencia y sus consecuencias en la dinámica económica es bastante reciente. Gabaix (2011) demuestra empíricamente que los shocks individuales que afectan a algunas “grandes” empresas influyen significativamente en las fluctuaciones del ciclo económico de Estados Unidos. Gabaix sostiene que si los shocks idiosincrásicos de unas pocas grandes empresas de un país contribuyen significativamente



a las variaciones del Producto Interior Bruto (PIB), su economía se puede definir como “economía granular”.

Desde una perspectiva económica general, la idea de la economía granular cuestiona profundamente el planteamiento macroeconómico convencional. En macroeconomía, el marco teórico establecido sostiene que los shocks idiosincrásicos de las empresas se cancelan mutuamente al agregarse (Lucas, 1977), contribuyendo de manera marginal a las fluctuaciones económicas. La intuición está basada en el “argumento de la diversificación” (*diversification argument*), que describe la economía como un conjunto de empresas pequeñas cuyos shocks son independientes. Son los shocks exógenos y sistémicos (es decir, que afectan la economía en su conjunto) los que generan las fluctuaciones del ciclo económico. En este marco, la economía se puede representar con el supuesto del agente representativo, es decir, un único consumidor y una única empresa. Esta hipótesis asume implícitamente un alto grado de homogeneidad entre el tamaño de las empresas de una economía. Sin embargo, si algunas empresas tienen un impacto significativo a nivel agregado, el argumento de la diversificación no es válido y, en consecuencia, el supuesto de la empresa representativa tampoco. Una economía granular se puede representar como el conjunto de una multitud de pequeñas unidades (pequeñas y medianas empresas), cuya evolución individual no tiene esencialmente ningún impacto a nivel agregado, y pocas unidades granulares (decenas de grandes empresas), cuyas fluctuaciones contribuyen significativamente a la variabilidad del ciclo económico. La visión de la economía como sistema granular plantea nuevas posibilidades para comprender la dinámica macroeconómica (Carvalho y Grassi, 2019) y novedosas implicaciones a la hora de aplicar políticas de estabilización del ciclo económico.

Blanco-Arroyo et al. (2018), siguiendo la metodología propuesta por Gabaix (2011), encuentran evidencia empírica que respalda la hipótesis granular en el caso de la economía española. El objetivo de este artículo es comprobar si la hipótesis granular es válida para subunidades económicas como son las comunidades autónomas. En particular, se evalúa si la variabilidad del ciclo económico de las comunidades con mayor impacto económico en España está relacionada con los shocks idiosincrásicos de un conjunto de grandes empresas. Las Comunidades escogidas para el análisis de acuerdo a



su relevancia sobre el PIB español y mayor número de grandes empresas son: la Comunidad de Madrid, Cataluña, la Comunidad Valenciana y el País Vasco. El objetivo final del análisis propuesto es determinar si la hipótesis granular es una aproximación válida para economías pequeñas con una baja diversificación sectorial.

Como primera aproximación a la hipótesis de granularidad de la economía española, la Figura 1 presenta las ventas de las 100 empresas más grandes de España por volumen de ventas y de la Comunidad de Madrid en relación al PIB de España. Como se puede observar, las empresas más grandes de España representan entorno al 25% del PIB durante los años objeto de estudio, y en Comunidad de Madrid entorno 20% de media. Basados en este primer análisis, se puede intuir que existe una gran heterogeneidad en el tamaño de las empresas de España, así como en la Comunidad de Madrid. La pregunta a plantear a continuación es si esta heterogeneidad en el tamaño de las empresas guarda relación con las fluctuaciones macroeconómicas.

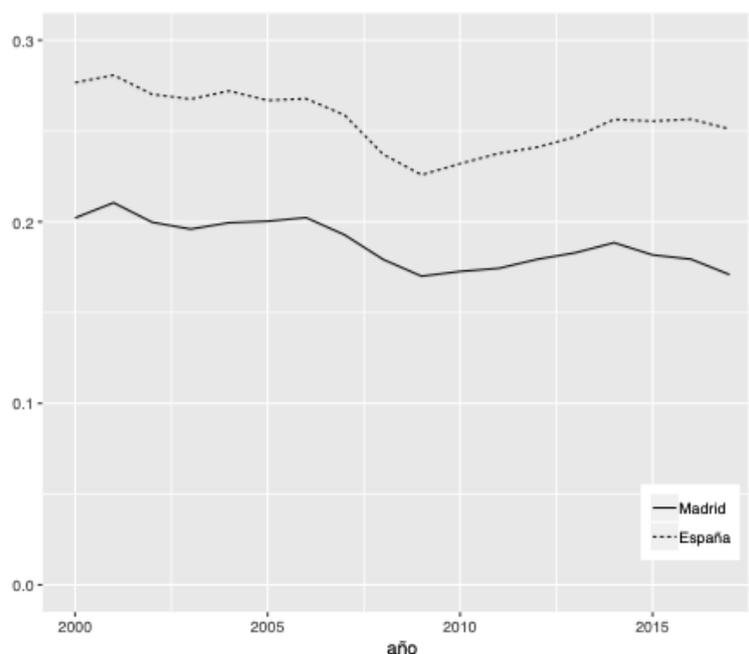


Figura 1. Total de ventas de las 100 empresas más grandes en España y en la Comunidad de Madrid en relación PIB español. No han sido consideradas las empresas cuya actividad principal esta relacionada con hidrocarburos, energía y finanzas.



Los resultados muestran que las empresas granulares constituyen una proporción marginal del número total de empresas en las regiones estudiadas. Sin embargo, encontramos que tanto en País Vasco como Comunidad Valenciana, una vez Mercadona es tratada individualmente, los shocks de productividad de estas empresas son responsables de una fracción significativa de las fluctuaciones en la tasa de crecimiento del PIB. Por el contrario, la Comunidad de Madrid y Cataluña, las dos comunidades con mayor aportación al PIB español y con mayor concentración de grandes empresas, no pueden ser consideradas como granulares.

El resto del trabajo está estructurado en los siguientes apartados. En el apartado 2 se expone el conjunto de datos utilizado y un resumen estadístico. El apartado 3 está destinado a la cuantificación de la heterogeneidad en el tamaño de las empresas que componen la muestra. En el apartado 4 se presenta el marco teórico, en cual se emplea el modelo de *Economy Islands* de Lucas para analizar si se cumple el argumento de la diversificación, y en contraposición se considera un modelo que tiene en cuenta la heterogeneidad en el tamaño de las empresas. En el apartado 5 se contrasta la hipótesis granular en cada una de las comunidades estudiadas. Por último, el apartado 6 presenta las conclusiones.

2. Datos y análisis descriptivo

Los datos utilizados en este estudio provienen de la base de datos Eurostat y Sistema de Análisis de Balances Ibéricos (SABI). Los datos macroeconómicos extraídos de Eurostat han sido el PIB real de España, PIB per cápita real y el deflactor del PIB. Además, se ha obtenido el PIB nominal y el PIB per cápita nominal de cada una de las comunidades y ciudades autónomas, los cuales han sido deflactados utilizando deflactor del PIB de España. Por otro lado, los datos microeconómicos utilizados para llevar a cabo el análisis han sido extraídos de la base de datos Sistemas de Análisis de Balances Ibéricos (SABI), elaborada por Bureau Van Dijk.



El periodo de estudio es de 17 años y está comprendido entre los años 2000 y 2017.¹ En este lapso de tiempo España ha experimentado un ciclo económico completo. La fase expansiva del mismo comprende el periodo 2001-2007 y la fase recesiva el periodo 2008-2013. Dado que nuestro análisis está centrado en las empresas más grandes, consideramos una muestra compuesta por aproximadamente las 50.000 empresas más grandes en SABI por volumen de ventas. Para cada una de las empresas disponemos de datos sobre volumen de ventas, número de empleados y código SIC, que identifica el sector de actividad de la empresa.² Cabe señalar que tanto las ventas como el número de empleados se refieren al territorio nacional. La falta de disponibilidad de datos a nivel regional es la principal limitación a la que nos enfrentamos en este trabajo.

Gabaix (2011) estudia las perturbaciones de productividad idiosincrásicas de las empresas. Las empresas cuyas perturbaciones pueden estar afectadas por factores exógenos no son incluidas en el análisis. Por esta razón, las empresas cuya actividad principal está relacionada con los hidrocarburos (códigos SIC 1311, 1389, 2911, 2999, 4922, 4923, 4924 and 5172), energía (códigos SIC entre 4900 y 4940), finanzas (códigos SIC entre 6000 y 6999) o administración pública (códigos SIC entre 9000 y 9799) son eliminadas de la muestra. La inclusión de estas empresas en el análisis puede llevar a distorsiones en los resultados debido a la naturaleza exógena de las fluctuaciones en sus ventas. Las empresas relacionadas con el sector de los hidrocarburos, por ejemplo, están fuertemente condicionadas por las fluctuaciones en el precio del petróleo. Estas empresas representan un 6.6% de la muestra.

SABI dispone de información sobre el domicilio social de las empresas, permitiéndonos poder localizar la comunidad autónoma a la que pertenece cada una. Sin embargo, las ventas y el número de empleados no están disponibles a nivel regional. Como se muestra en la Figura 2a y 2b, las empresas más grandes se concentran en unas pocas regiones con una alta contribución al PIB nacional. La Figura 2a muestra la contribución media al PIB español por comunidad autónoma en el periodo 2000-2017. La contribución es el PIB real

¹ El periodo de estudio está condicionado a la disponibilidad de datos macroeconómicos regionales.

² *Standard Industrial Classification* (SIC) es un sistema de clasificación industrial basado en código de cuatro dígitos.



de la comunidad como fracción del PIB real de España. En promedio, Cataluña es la comunidad que más aporta (19.8%), seguida de la Comunidad de Madrid (19.3%), Andalucía (14.2%), la Comunidad Valenciana (9.9%) y el País Vasco (6.4%). Por otra parte, la Figura 2b muestra el peso medio de las 100 empresas más grandes de cada comunidad autónoma en el periodo 2000-2017. El peso es la suma de las ventas reales que las 100 empresas más grandes de la comunidad tienen en España como fracción del PIB real español. la Comunidad de Madrid (18.8%) es la comunidad donde mayor tamaño relativo tienen las empresas más grandes, seguida de Cataluña (6.7%), Galicia (3.2%), la Comunidad Valenciana (3.1%) y el País Vasco (2.7%). Utilizando estos datos como referencia, decidimos centrar nuestro análisis en aquellas comunidades que presentan la mayor combinación de impacto sobre el PIB de España y peso de las grandes empresas: Comunidad de Madrid, Cataluña, Comunidad Valenciana y País Vasco. Andalucía y Galicia sólo presentan un alto valor en uno de los criterios, por eso no han sido incluidas.

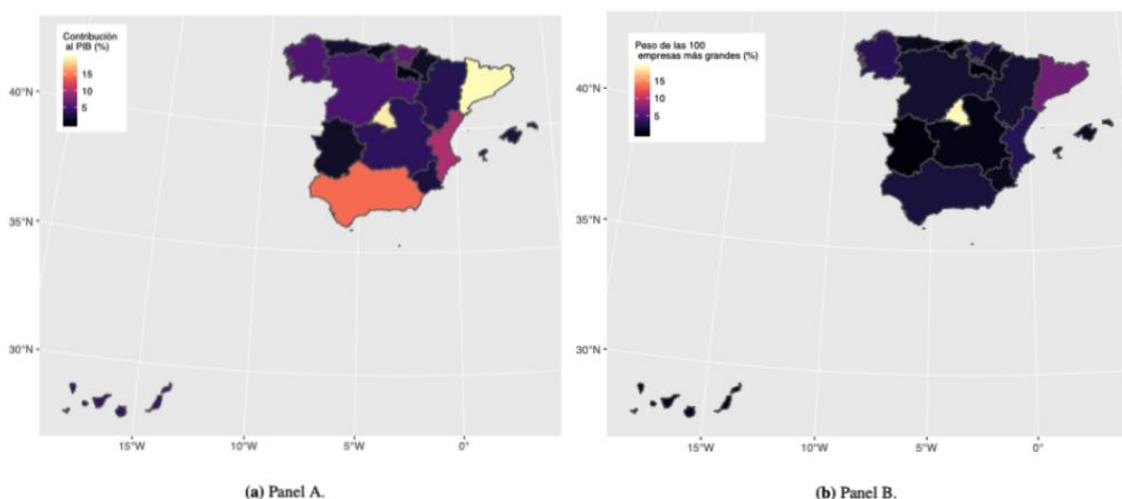


Figura 2. Panel A: contribución media al PIB español por comunidad autónoma en el periodo 2000- 2017. La contribución de cada comunidad autónoma es la ratio entre PIB real de la comunidad y PIB real español. Panel B: peso medio de las 100 empresas más grandes por comunidad autónoma en el periodo 2000-2017. El peso es la suma de las ventas de las 100 empresas más grandes como fracción del PIB real español.

El número de empresas que componen nuestra muestra después del proceso de filtrado es 46.307, de las cuales 21.4% tienen su domicilio social en la Comunidad de Madrid, 22% en Cataluña, 10.3% en la Comunidad Valenciana y 6.5% en el País Vasco. Por lo tanto, más del 60% de la muestra se encuentra domiciliada en estas comunidades. Este número



es representativo de la distribución geográfica de las empresas en España, ya que el 50% de su totalidad se encuentran domiciliadas en estas cuatro regiones.³ La Tabla 1 presenta un resumen estadístico de las ventas de las empresas que componen la muestra y las submuestras que representan las comunidades. Observamos que el crecimiento de las ventas agregadas – suma de las ventas de las empresas de todas las empresas – es ligeramente inferior al crecimiento del PIB tanto en la muestra como en cada una de las regiones, excepto la Comunidad Valenciana, donde el crecimiento de las ventas en el periodo de estudio son casi el doble que el crecimiento del PIB. En lo que respecta a la variabilidad de las ventas, observamos que tanto la muestra agregada como la regional tienen unas fluctuaciones mayores que el PIB. Además, la correlación entre el crecimiento de las ventas agregadas y del PIB es alta, sólo Cataluña está por debajo del 50%. En la tabla también se muestra el peso – suma de las ventas como fracción del PIB de España – medio por quintil. Observamos que existe una gran heterogeneidad en la distribución del tamaño de las empresas en España y las comunidades. Las empresas más grandes representan una gran parte de la actividad macroeconómica. En la sección 3 estudiaremos en detalle la distribución del tamaño de las empresas medida por las ventas, con el objetivo de identificar la distribución teórica que mejor la describe y así poder analizar si las presentes características pueden generar el comportamiento granular en la economía.

Tabla 1. Resumen estadístico. Las ventas agregadas se refieren a la suma de las ventas reales de todas las empresas. La variable peso indica la suma de las ventas como fracción del PIB real de España. España se

³ Calculado a partir de los datos de Eurostat sobre demografía empresarial a nivel regional en el año 2016.



refiere a la muestra sin considerar el domicilio social de las empresas que la forman. Las comunidades son submuestras en las que se ha considerado el domicilio social. Periodo de tiempo: 2000-2017.

	España	Madrid	Cataluña	País Vasco	Comunidad Valenciana
Número empresas	46307	9917	10165	3012	4784
Media					
Crecimiento PIB	0.017				
Crecimiento ventas agregadas	0.016	0.011	0.0069	0.017	0.034
Desviación estándar					
Crecimiento PIB	0.026				
Crecimiento ventas agregadas	0.046	0.049	0.051	0.064	0.051
Correlación crecimiento PIB y crecimiento ventas agregadas	0.65	0.60	0.49	0.66	0.67
Peso medio					
Top 10	0.087	0.072	0.027	0.0096	0.018
Top 100	0.26	0.19	0.067	0.027	0.028
0-20 percentil	0.77	0.33	0.14	0.039	0.040
21-40 percentil	0.099	0.029	0.022	0.0071	0.0093
41-60 percentil	0.054	0.014	0.012	0.0039	0.0053
61-80 percentil	0.033	0.0075	0.0077	0.0024	0.0033
81-100 percentil	0.015	0.0030	0.0036	0.0011	0.0015

La Tabla 2 recoge el peso medio por sector en España y en el conjunto que forman las comunidades estudiadas. Las empresas de la muestra pertenecientes al sector de ventas al por mayor, tanto de bienes duraderos como no duraderos, tienen un mayor peso en la economía; seguido del sector de equipo de transporte, alimentos y productos afines y tiendas de comida. En el conjunto de las comunidades ventas al por mayor es también el sector cuyas empresas presentan el mayor peso, seguido de tiendas de comida, servicios para empresas y comunicaciones.

3. Distribución del tamaño de las empresas

Como se observa en la Tabla 1, existe una gran heterogeneidad en el tamaño de las empresas. El objetivo de esta sección es identificar el tipo de distribución subyacente. La Figura 3a muestra la Función de Distribución Acumulada (FDA) (x) de las ventas de las empresas en la muestra. Destaca que la distribución está altamente sesgada hacia la



derecha, lo que significa que mientras la mayor parte de empresas son de tamaño significativamente inferior a la media, hay un reducido número de empresas cuyo volumen de ventas es mucho más alto que el valor típico.

Además, si representamos la FDA con los ejes en escala logarítmica (Figura 3b) observamos que la cola superior de la distribución, aquella donde se encuentran las empresas de mayor tamaño, puede ser aproximada por una línea recta. Así, para $x \geq x_{min}$ tenemos $\ln P(x) = constante + \alpha \ln x$, donde α es una constante independiente de las unidades en la que se expresan los datos (Gabaix, 2009) y representa el índice de la cola o parámetro de escala. Si la distribución presenta este comportamiento se dice que sigue una ley de potencia (*power law*). Esta familia de distribuciones se caracteriza por ser de “cola gruesa”, es decir, tiene una cola que decae a cero más lentamente que una función exponencial. Este hecho hace que las perturbaciones a las empresas más grandes, situadas en la cola de la distribución, no desaparezcan a nivel agregado (Gabaix, 2011).

Hill (1975) estima el parámetro de escala α por máxima verosimilitud y obtiene

$$\hat{\alpha} = 1 + n \left(\sum_{i=1}^n \ln \frac{x_i}{x_{min}} \right)^{-1} \quad (1)$$



Tabla 2. Peso medio en porcentaje por sector en la muestra de España y en el conjunto que forman las cuatro comunidades objeto de estudio: la Comunidad de Madrid, Cataluña, el País Vasco y la Comunidad Valenciana. La variable peso se refiere a la suma de las ventas reales como fracción del PIB real de España. Periodo de tiempo: 2000-2017.

Division	Sector	Peso		Sector	Peso			
		España	Comunidades		España	Comunidades		
Agricultura, silvicultura y pesca	Producción agrícola - cultivos	01	0.19	0.06	42	1.18	0.98	
	Producción agrícola - ganadería	02	0.62	0.45	43	0.20	0.27	
	Servicios agrícolas	07	0.07	0.04	44	1.22	1.51	
	Silvicultura	08	0.02	0.01	45	1.11	1.32	
	Pesca, caza y trampa	09	0.06	0.04	46	0.11	0.15	
	Minería	Minería de metales	10	0.05	0.03	47	2.02	1.75
		Minería del carbón	12	0.09	0.04	48	3.69	5.07
		Minerales no metálicos, excepto combustibles	14	0.21	0.19	49	1.05	1.10
	Construcción	Contratistas generales de la construcción	15	3.41	3.42	50	12.21	13.49
Construcción pesada, excepto edificio		16	2.09	2.46	51	12.85	11.15	
Contratistas comerciales especiales		17	1.50	1.65	52	0.11	0.11	
Producción	Alimentos y productos afines	20	6.45	4.73	53	1.43	2.03	
	Productos del tabaco	21	0.30	0.44	54	5.36	6.51	
	Productos textiles de molino	22	0.33	0.42	55	0.78	0.76	
	Prendas de vestir y otros productos textiles	23	0.34	0.30	56	0.84	0.63	
	Madera y productos de madera	24	0.40	0.26	57	0.46	0.49	
	Muebles y accesorios	25	0.24	0.20	58	0.46	0.58	
	Papel y productos afines	26	1.29	1.28	59	0.92	1.10	
	Impresión y publicación	27	0.85	1.06	70	0.90	0.52	
	Productos químicos y afines	28	4.18	5.21	72	0.12	0.15	
	Productos de caucho y plásticos, misceláneos	30	1.06	1.24	73	4.55	6.02	
	Cuero y productos de cuero	31	0.08	0.08	75	0.43	0.48	
	Productos de plástico, arcilla y vidrio	32	1.62	1.84	76	0.10	0.07	
	Industrias de metales primarios	33	2.55	2.15	78	0.26	0.35	
	Productos metálicos fabricados	34	1.55	1.37	79	0.79	0.84	
	Maquinaria y equipo industrial	35	1.30	1.28	80	0.59	0.56	
	Equipos electrónicos y otros equipos eléctricos	36	1.75	1.84	81	0.11	0.15	
	Equipos de transporte	37	6.46	4.97	82	0.17	0.21	
	Instrumentos y productos relacionados	38	0.25	0.35	83	0.36	0.50	
	Industrias manufactureras, misceláneos	39	0.14	0.16	84	0.02	0.02	
	Transporte y servicios públicos	Transporte ferroviario	40	0.13	0.19	86	0.01	0.01
		Tránsito de pasajeros local e interurbano	41	0.76	0.84	87	2.11	2.55

regiones estudiadas el parámetro de escala es superior a 3, indicando que la economía de las regiones analizadas puede ser granular.

Tabla 3. Parámetros estimados de la distribución del tamaño de las empresas. Los valores presentados son la media en el tiempo. n_{cola} indica el número de empresas cuya distribución del tamaño puede ser descrita por una power law. \hat{x}_{min} son las ventas mínimas estimadas a partir de las cuales se puede observar el comportamiento power law. $\hat{\alpha}$ es el parámetro de escala de la power law. \hat{m} y $\hat{\sigma}^2$ representan la media y la varianza en la distribución log-normal, y están calculadas dado \hat{x}_{min} para poder comparar el ajuste de la distribución power law y log-normal. \mathcal{R} es la ratio de probabilidades entre las dos distribuciones consideradas como posible ajuste de los datos. p-valor indica si \mathcal{R} es significativamente distinto de cero.

	n_{cola}	\hat{x}_{min}	$\hat{\alpha}$	\hat{m}	$\hat{\sigma}^2$	\mathcal{R}	p-valor
España	4094	$4.27 \cdot 10^7$	2.14	-12.77	5.17	-1.07	0.32
Madrid	1500	$3.52 \cdot 10^7$	1.91	9.07	3.37	-2.03*	0.06
Cataluña	826	$4.52 \cdot 10^7$	2.25	-10.21	4.77	-0.46	0.65
País Vasco	424	$3.71 \cdot 10^7$	2.23	7.39	2.87	-1.01	0.38
Comunidad Valenciana	659	$3.34 \cdot 10^7$	2.45	-	-	-	-
Comunidad Valenciana excluyendo Mercadona	706	$2.99 \cdot 10^7$	2.44	-35.19	3.20	-1.54	0.27

España es la muestra sin distinguir las empresas por la región donde están domiciliadas.
 En Comunidad Valenciana la distribución del tamaño de las empresas presenta una heterogeneidad extrema, debido a Mercadona, que provoca que los parámetros de la distribución log-normal no puedan ser estimados con fiabilidad. Por eso decidimos reestimar estos parámetros excluyendo esta empresa.

* Estadísticamente significativo al 10%.

La distribución *power law*, en numerosos casos empíricos, no puede estadísticamente distinguirse de la distribución log-normal. Por este motivo, decidimos contrastar cuál de las dos distribuciones se ajusta mejor a la distribución del tamaño de las empresas. Para poder comparar ambas distribuciones estimamos los parámetros que caracterizan la log-normal, media y varianza, dado \hat{x}_{min} y calculamos la ratio de probabilidades \mathcal{R} (Clauset et al., 2009), la cuál puede tener valores negativos, positivos o ser cero. Un valor negativo nos indica que la distribución log-normal resulta ser más adecuada, mientras que si es positivo indica que la distribución la *power law* es mejor. Si ambas distribuciones no pueden ser diferenciadas, entonces \mathcal{R} es cero. En la Tabla 3 se observa que en todos los casos el valor de \mathcal{R} es negativo, sugiriendo que la distribución log-normal describe mejor la distribución del tamaño de las empresas que la distribución *power law*. Visualmente también se puede apreciar que el ajuste de la log-normal es mejor (ver figuras 3b y 4).



verosimilitud de la distribución power-law y log-normal. Para la Comunidad Valenciana se ha calculado el ajuste de la distribución excluyendo las empresas Mercadona (figura insertada en (d)).

4. Marco teórico

4.1 Empresas independientes

Con el fin de ilustrar el planteamiento propuesto por Gabaix (2011), recurrimos a un modelo basado en el marco teórico de *Economy Islands* de Lucas (1972, 1973, 1975). Este modelo asume que todas las empresas producen bienes finales y carecen de vínculo entre ellas respecto al insumo utilizado en la producción, es decir, las empresas no son proveedoras de otras. Bajo esta hipótesis el PIB en el tiempo t es $Y_t = \sum_{i=1}^N S_{it}$, donde S_{it} es la cantidad producida y vendida – nos referiremos a esta cantidad como ventas – de la empresa i durante el tiempo t . Por lo tanto, el crecimiento del PIB de la economía es:

$$\frac{\Delta Y_t}{Y_{t-1}} = \sum_{i=1}^N \frac{S_{i,t-1}}{Y_{t-1}} \frac{\Delta S_{it}}{S_{i,t-1}} \quad (2)$$

donde $\Delta Y_t = Y_t - Y_{t-1}$ y $\Delta S_{it} = S_{it} - S_{i,t-1}$ son las variaciones en un intervalo de tiempo del PIB y las ventas de la empresa i , respectivamente. N representa el número de empresas independientes (Islas) que constituyen la economía. La ecuación (2) indica que el crecimiento del PIB es la suma del crecimiento ponderado de cada empresa con pesos $S_{i,t-1}/Y_{t-1}$. Dado que las empresas son independientes, sus tasas de crecimiento se deben a perturbaciones idiosincrásicas, por ejemplo, huelgas de los trabajadores, cambio en la directiva de la empresa, etc.

Siguiendo el planteamiento de Gabaix (2011), se puede definir la tasa de crecimiento de la empresa i como $\Delta S_{it}/S_{i,t-1} = \sigma_i \varepsilon_{i,t+1}$, donde σ_i es la desviación típica en el crecimiento de las ventas de la empresa i y $\varepsilon_{i,t+1}$ es una perturbación aleatoria con media cero y varianza unitaria, y no correlacionada con las perturbaciones de otras empresas. Si, además, se asume que todas las empresas que forman esta economía tienen aproximadamente la misma volatilidad ($\sigma_i \equiv \sigma$), la relación entre la volatilidad del PIB y la volatilidad de las empresas puede ser expresada como:



$$\sigma_{PIB}^h = \sigma \sqrt{\sum_{i=1}^N \left(\frac{S_{it}}{Y_t}\right)^2} = \sigma h \quad (3)$$

La ecuación (3) implica que la volatilidad de la tasa de crecimiento del PIB σ_{PIB} es directamente proporcional a la volatilidad del crecimiento de las empresas con un coeficiente que depende de la distribución del tamaño de las empresas (h). El coeficiente h es la raíz cuadrada del índice Herfindahl-Hirschman (HH) de las ventas. Un índice de concentración que asigna un peso mayor a las empresas con mayor tamaño. El valor de este índice depende directamente de la distribución del tamaño relativo de las empresas (Naldi, 2003).

Asumiendo que todas las empresas tienen el mismo tamaño, tal y como se asume en los modelos macroeconómicos convencionales basados en la empresa representativa, la volatilidad del PIB pasaría a ser:

$$\sigma_{PIB}^I = \frac{\sigma_S}{\sqrt{N}} \quad (4)$$

donde σ_S indica la desviación típica en el modelo de empresas independientes.

Con el fin de comprobar si este es el caso de las comunidades objeto de estudio, hemos estimado la volatilidad del crecimiento de las ventas de las 100 empresas más grandes de cada comunidad, así como h . La Tabla 4 presenta la volatilidad media del crecimiento de las ventas σ_S y el índice HH medio h . Además, N es el número total de empresas en cada una de las regiones en el año 2016 según Eurostat. Por último, se ha incluido también la volatilidad del PIB como medida de referencia para comparar la estimación. Según la ecuación (4), σ_{PIB}^I es 0.013% en Comunidad de Madrid, 0.007% en Cataluña, 0.019% en País Vasco y 0.012% en Comunidad Valenciana. El resultado obtenido para cada una de las comunidades no es consistente con la volatilidad observada del PIB, que resulta ser de un orden de magnitud mayor.

Este simple cálculo indica la razón por la cual los modelos macroeconómicos convencionales (por ejemplo, Kydland y Prescott (1982) y King et al. (1988)), basados en el paradigma de la empresa representativa, consideran las perturbaciones agregadas y



Productividad Total de los Factores (PTF) puede ser explicada a partir de la suma ponderada de los crecimientos de productividad:⁶⁷

$$\frac{dPTF}{PTF} = \sum_i \frac{S_i}{Y} d\pi_i \quad (5)$$

donde $d\pi_i$ es el crecimiento de la productividad y los pesos son S_i/Y , también conocidos como pesos de Domar (1961). Estos pesos, de acuerdo con el teorema de Hulten, son un estadístico suficiente para calcular las perturbaciones de la PTF, por lo que no es necesario conocer la matriz input-output.

La correlación directa entre productividad y actividad económica es la base de numerosos modelos macroeconómicos de crecimiento. Por consiguiente,

$$\frac{dY}{Y} = \mu \sum_i \frac{S_i}{Y} d\pi_i \quad (6)$$

donde μ es el factor de uso. Calculando la volatilidad del crecimiento del PIB se obtiene:

$$\sigma_{PIB}^h = \mu \sigma_{PTF} = \mu \sigma_{\pi} h \quad (7)$$

donde $\sigma_{\pi} = \sqrt{\text{var}(d\pi_i)}$ y σ_{PTF} es la desviación típica de la productividad total de los factores.

Con el objetivo de verificar si este planteamiento se aproxima más a la volatilidad observada del crecimiento del PIB, utilizamos el valor de $\mu = 2.6$ propuesto por Gabaix (2011). Bajo el supuesto de que todas las comunidades tienen el mismo uso de los factores, $\hat{\sigma}_{PIB}$ es 1.9% en Comunidad de Madrid, 1.2% en Cataluña, 3.3% en País Vasco y 5.1% en Comunidad Valenciana. Estos valores son los que más se aproximan a la volatilidad observada del PIB (Tabla 4). Este simple modelo que tiene en cuenta las

⁶ En el modelo de las islas se considera el crecimiento de las ventas, mientras que en el teorema de Hulten se relaciona el crecimiento de la productividad de las empresas con las variaciones de la PTF.

⁷ La productividad del trabajo se ha medido a través de las ventas por empleado. La ratio ventas por empleado tiene la ventaja de ser fácilmente calculada con los datos disponibles, y es la medida utilizada por Gabaix en el trabajo original. Por su parte, la volatilidad es la raíz cuadrada de la varianza transversal en el año t .



interacciones entre las empresas, junto a la gran heterogeneidad en sus tamaños, llega a predecir el orden de magnitud de las fluctuaciones del PIB de las comunidades consideradas.

Dado que un pequeño número de grandes empresas representa una fracción significativa del PIB, como se muestra en la Tabla 1, y considerar su heterogeneidad en el tamaño mejora enormemente la volatilidad predicha del crecimiento PIB, podemos preguntarnos: ¿es posible que los shocks idiosincrásicos de las grandes empresas representen una fracción significativa de las fluctuaciones del ciclo económico?

5. Comprobación empírica de la hipótesis granular

En el apartado anterior hemos calibrado el modelo basándonos en el teorema de Hulten, en el cual las empresas están vinculadas a través del mercado de insumos. Teniendo en cuenta la heterogeneidad en el tamaño de las empresas, el modelo es capaz de predecir una magnitud de las fluctuaciones en la actividad económica próxima a la observada. Con el fin de analizar si la hipótesis de la economía granular se cumple para las comunidades objeto de estudio, utilizamos la metodología propuesta por Gabaix (2011). Este autor propone una medida de los shocks idiosincrásicos llamada “residuo granular”, basada en la ecuación (5) derivada desde el teorema de Hulten. El residuo granular es una medida que recoge los shocks idiosincrásicos de las K empresas más grandes, ponderadas por su peso en la economía:

$$\Gamma_t = \sum_{i=1}^K \frac{S_{i,t-1}}{Y_{t-1}} \varepsilon_{it} \quad (8)$$



donde ε_{it} es la perturbación a la empresa i en el tiempo t . Debido a que ε_{it} no es observado, debe ser estimado. Gabaix (2011) propone calcular el crecimiento de la productividad y sustraer el crecimiento medio.⁸⁹ El crecimiento de la productividad es

$$g_{it} = \Delta \ln \frac{S_{it}}{E_{it}} \quad (9)$$

donde E_{it} es el número de empleados en la empresa i en el tiempo t . La media es $\bar{g}_t = Q^{-1} \sum_{i=1}^Q g_{it}$, siendo $Q \geq K$. Por lo tanto, el residuo granular que calculamos esta determinado por

$$\Gamma_t = \sum_{i=1}^K \frac{S_{i,t-1}}{Y_{t-1}} (g_{it} - \bar{g}_t) \quad (10)$$

La idea básica es que las fluctuaciones idiosincrásicas de pocas empresas, $Q = K = 100$, sean suficientes para describir una fracción significativa de las variaciones del ciclo económico. Para cuantificar el poder explicativo regresamos por Mínimos Cuadrados Ordinarios (MCO) el crecimiento del PIB per cápita real (y_t) sobre el residuo granular (10), esto es

$$\frac{\Delta y_t}{y_{t-1}} = \text{constante} + \mu \Gamma_t + u_t \quad (11)$$

Donde $\Delta y_t = y_t - y_{t-1}$. Si la economía fuera granular, un número reducido de empresas, como son las 100 empresas más grandes, debería ser suficiente para explicar una fracción significativa de las fluctuaciones de PIB.

La principal limitación a la que tenemos que enfrentarnos es la falta de datos a nivel regional en SABI. Desafortunadamente, no disponemos de las cifras de ventas ni del número de empleados a nivel regional. Por lo que tenemos que estimar la fracción de

⁸ Cabe mencionar que Gabaix propone una estimación alternativa de ε_{it} basada en sustraer la media del crecimiento de la productividad del sector al que pertenece i . Debido al pequeño tamaño de nuestra muestra, el número de empresas por sector no es lo suficientemente grande para ser representativo del crecimiento medio de la productividad.

⁹ Para reducir el impacto de valores atípicos, atenúamos las perturbaciones estimadas más pequeñas que la probabilidad del 10% y más grandes que la probabilidad del 90%.



estas variables que la empresa i utiliza para la producción de los bienes vendidos en la comunidad. Es decir, calibramos la fracción de ventas en la comunidad y el número de empleados necesarios para producir estos bienes. Sabemos que tanto las ventas como el número de empleados de la empresa i en el tiempo t en la región donde esta domiciliada son una fracción $0 < \lambda \leq 1$ de las ventas y número de empleados que tiene en España, esto es $S_{it} = \lambda_{it} S_{it}$ y $E_{it} = \lambda_{it} E_{it}$, respectivamente. Por simplicidad, asumimos que esta fracción es la misma para todas las empresas y en el tiempo, así que $S_{it} = \lambda S_{it}$ y $E_{it} = \lambda E_{it}$.¹⁰ Dado que λ forma parte de S_{it} y E_{it} , la productividad de la empresa i , así como su crecimiento (ecuación (9)), en la región donde está domiciliada es la misma que tiene a nivel nacional, $G_{it} = g_{it}$. Sin embargo, los pesos en la ecuación (10) sí se ven afectados, puesto que las ventas en España son superiores a las de la región. Por consiguiente, el residuo granular en la región r sería

$$\Gamma_{rt} = \sum_{i=1}^K \frac{\lambda S_{i,t-1}}{Y_{r,t-1}} (g_{it} - \bar{g}_t) \quad (12)$$

Si estimamos el modelo (11), sin considerar el impacto de la fracción λ , la constante de proporcionalidad es $\hat{\gamma} = \mu\lambda$, donde μ representa el factor de uso regional. Para estimar λ , y hallar así μ , recurrimos a la ecuación (7). Manteniendo el supuesto mencionado anteriormente que asume que la productividad de la empresa i coincide con la productividad de la región donde está domiciliada y que h no se ve afectado si se asume que la fracción de las ventas en la región es para todas las empresas igual, tenemos que el factor de uso para cada comunidad esta determinado por

$$\mu = \frac{\sigma_{PIB}}{\sigma_{\pi} h} \quad (13)$$

¹⁰ Es importante tener en cuenta que la aproximación del coeficiente λ como una constante, asumiendo homogeneidad en el tiempo y entre las empresas, es muy restrictiva. En futuras investigaciones, incorporaremos la heterogeneidad de las empresas en nuestro análisis utilizando datos detallados a nivel regional, los cuales no están disponibles actualmente.

Tabla 6. Poder explicativo de las 100 empresas más grandes en cada comunidad una vez se ha tenido en cuenta la fracción de las ventas que las empresas tienen en la región donde están domiciliadas λ .

	Madrid	Cataluña	País Vasco	Comunidad Valenciana	Comunidad Valenciana 1
Constante	0.010* (0.005)	0.010 (0.008)	0.021*** (0.005)	0.005 (0.007)	0.008 (0.007)
Γ_t	3.005 (3.031)	5.361 (4.569)	1.895*** (0.240)	-1.569 (3.314)	5.64** (2.108)
Observaciones	17	17	17	17	17
R ²	0.043	0.058	0.481	0.026	0.151
R ² ajustado	-0.021	-0.005	0.446	-0.039	0.095

Errores robustos a autocorrelación en paréntesis.

* nivel de significatividad al 10%, ** 5% y *** 1%.

El caso de Comunidad Valenciana es particular porque el resultado está condicionado por una sola empresa: Mercadona. De acuerdo con las estimaciones realizadas por el Instituto Valenciano de Investigaciones Económicas (Ivie) (Ivie, 2018), la aportación de Mercadona al PIB español es actualmente 1.9%, y está en línea con la fracción del PIB español que representan las ventas de Mercadona (1.8%) en nuestra muestra. Sin embargo, dada la limitación que presentan los datos, cuando calculamos la fracción del PIB de Comunidad Valenciana que representan las ventas de Mercadona encontramos que es desproporcionada (19%). Para atenuar el impacto de Mercadona asumimos que las ventas y el número de empleados de Mercadona en la Comunidad Valenciana está determinada por el número de establecimientos en esta región. De acuerdo con la Memoria anual de Mercadona del año 2018, hay 1.636 establecimientos en España, de los cuales 296 están situados en Comunidad Valenciana,¹¹ representando así el 18%. Esta modificación en la distribución del tamaño de las empresas disminuye el grado de concentración, $h = 0.057$, y aumenta $\hat{\mu} = 4.95$. Para calcular el residuo granular asumimos que el resto de empresas son locales, por lo que $\lambda = 1$. La última columna de la Tabla 6 presenta los resultados de esta estimación. Cuando Mercadona es tratada individualmente, la economía de Comunidad Valenciana es granular, ya que las

¹¹ El número de establecimientos por provincia ha sido extraído del listado provincial de Mercadona.



perturbaciones idiosincrásicas de las empresas más grandes explican una fracción significativa de las fluctuaciones agregadas.

6. Conclusiones

Este trabajo investiga la existencia de comportamiento granular a nivel regional, esto es, si las perturbaciones idiosincrásicas a las empresas más grandes de una región explican una fracción significativa de las fluctuaciones agregadas de dicha región. Utilizando datos de España, construimos el residuo granular, propuesto por Gabaix (2011), para aquellas comunidades autónomas con mayor contribución al PIB español y mayor número de grandes empresas. Estas comunidades son: la Comunidad de Madrid, Cataluña, la Comunidad Valenciana y el País Vasco. Dado que no disponemos de datos microeconómicos a nivel regional, proponemos un método de estimación que permite aproximar el impacto que las empresas tienen en la región donde están domiciliadas.

Encontramos que País Vasco y Comunidad Valenciana, una vez Mercadona es tratada individualmente, son economías granulares. En el primer caso las perturbaciones a las 100 empresas más grandes explican casi la mitad de las fluctuaciones agregadas, mientras que en el segundo explican aproximadamente el 10%. Comunidad de Madrid y Cataluña, las dos comunidades con mayor aportación al PIB español y con mayor concentración de grandes empresas, no pueden ser consideradas como granulares.

El trabajo futuro debe centrarse en desarrollar una mejor estimación de la fracción que las empresas nacionales dedica a la producción de bienes destinados a la región donde se encuentran domiciliadas. Esto permitirá establecer con mayor precisión el impacto económico de las empresas a nivel regional.

7. Agradecimientos

Agradecemos la financiación de la Universitat Jaume I, España, en el marco del proyecto UJI-B2018-77, a la Generalitat Valenciana en el marco del proyecto AICO/2018/036 y al Ministerio de Ciencia, Innovación y Universidades en el marco del proyecto RTI2018-096927-B-I00. Omar Blanco-Arroyo reconoce el apoyo financiero de la Universitat



Jaume I, España (PREDOC/2017/55). David Vidal-Tomás y Alba Ruiz-Buforn reconocen el apoyo financiero del Ministerio de Educación de España, FPU2015/01434 y FPU2014/01104, respectivamente.

Referencias

- Blanco-Arroyo, O., Ruiz-Buforn, A., Vidal-Tomás, D. and Alfarano, S. (2018): '*On the determination of the granular size of the economy*', *Economics Letters* 173, 35–38.
- Carvalho, V. M. and Grassi, B. (2019): '*Large firm dynamics and the business cycle*', *American Economic Review* 109(4), 1375–1425.
- Clauset, A., Shalizi, C. R. and Newman, M. E. (2009): '*Power-law distributions in empirical data*', *SIAM review* 51(4), 661–703.
- Domar, E. D. (1961): '*On the measurement of technological change*', *The Economic Journal* 71(284), 709–729.
- Gabaix, X. (2009): '*Power laws in economics and finance*', *Annu. Rev. Econ.* 1(1), 255–294.
- Gabaix, X. (2011): '*The granular origins of aggregate fluctuations*', *Econometrica* 79, 733–772.
- Gabaix, X. and Ibragimov, R. (2011): '*Rank-1/2: a simple way to improve the ols estimation of tail exponents*', *Journal of Business & Economic Statistics* 29(1), 24–39.
- Hill, B. M. (1975): '*A simple general approach to inference about the tail of a distribution*', *The annals of statistics* 3(5), 1163–1174.
- Hulten, C. R. (1978): '*Growth accounting with intermediate inputs*', *The Review of Economic Studies* 45(3), 511–518.
- Ivie (2018): '*El impacto económico de Mercadona*', <https://info.mercadona.es/document/es/resumen-ejecutivo-sobre-el-impacto-economico-de-mercadona-2018-ivie.pdf>.
- King, R. G., Plosser, C. I. and Rebelo, S. T. (1988): '*Production, growth and business cycles: I. The basic neoclassical model*', *Journal of monetary Economics* 21(2-3), 195–232.
- Kydland, F. E. and Prescott, E. C. (1982): '*Time to build and aggregate fluctuations*', *Econometrica: Journal of the Econometric Society* pp. 1345–1370.



- Lucas, R. E. (1972), 'Expectations and the neutrality of money', *Journal of economic theory* 4(2), 103–124.
- Lucas, R. E. (1973): 'Some international evidence on output-inflation tradeoffs', *The American Economic Review* 63(3), 326–334.
- Lucas, R. E. (1975): 'An equilibrium model of the business cycle', *The Journal of Political Economy* pp. 1113–1144.
- Lucas, R. E. (1977): *Understanding business cycles*, in 'Carnegie-Rochester conference series on public policy', Vol. 5, North-Holland, pp. 7–29.
- Lux, T. and Alfarano, S. (2016): 'Financial power laws: Empirical evidence, models, and mechanisms', *Chaos, Solitons & Fractals* 88, 3–18.
- Lux, T. and Marchesi, M. (1999): 'Scaling and criticality in a stochastic multi-agent model of a financial market', *Nature* 397(6719), 498.
- Mantegna, R. N. and Stanley, H. E. (1999): *Introduction to econophysics: correlations and complexity in finance*, Cambridge university press.
- Naldi, M. (2003): 'Concentration indices and zipf's law', *Economics Letters* 78(3), 329–334.
- Pareto, V. (1896): 'Cours d' économie politique reprinted as a volume of oeuvres compl'etes', Droz, Geneva 1965.
- Stanley, H. E. (1971): *Phase transitions and critical phenomena*, Clarendon Press, Oxford.
- Vuong, Q. H. (1989): 'Likelihood ratio tests for model selection and non-nested hypotheses', *Econometrica: Journal of the Econometric Society* pp. 307–333.

