estructuraeconòmica
ECONOMÍA APLICADA II

The granularity of Spanish exports

by

Juan de Lucio
Raúl Mínguez
Asier Minondo
and
Francisco Requena

January 2017

Working Papers in Applied Economics

WPAE-2017-01

ECONOMÍA APLICADA II

VNIVERSITAT DE VALÈNCIA

Departamento de Estructura Económica (Economía Aplicada II), Facultad de Economía, Universitat de València
Avda dels Tarongers, s/n, 46022 – Valencia, Spain
Phone +34 96 382 83 49 / Fax +34 96 382 83 54
http://www.estructuraeconomica.es
e-mail: investigacion.estructura@uv.es

ISSN 2172-3036

The granularity of Spanish exports*

Juan de Lucio† Raúl Mínguez‡ Asier Minondo§
Francisco Requena¶

December 21, 2016

Abstract

Using data for the universe of exporters, we show that few firms dominate exports in Spain. For example, in 2015 half of Spanish exports were accounted by the top 200 firms. This concentration has not changed substantially over the period 1997-2015. The dominance of few firms, a phenomenon denoted as granularity, is larger in exports than in sales or employment. Granularity also defines the specialization of Spanish exports. If top exporters disappeared, Spain would lose revealed comparative advantage in industries accounting for 26% of Spanish exports. Finally, granularity explains around one-third of the fluctuation of Spanish exports.

JEL: F1, F10, F23

Keywords: exports, granularity, superstars, Spain, firm-level data.

*We thank Francisco Olarte for preparing the SABI data used in the empirical analyses and Jordi Paniagua for valuable suggestions and comments. We also thank the Customs and Excise Department of the Spanish Revenue Agency (AEAT) for providing essential information for this paper. We gratefully acknowledge financial support from the Spanish Ministry of Economy and Competitiveness (MINECO ECO2015-68057-R and ECO2016-79650-P, co-financed with FEDER), the Basque Government Department of Education, Language Policy and Culture (IT885-16), and the Generalitat Valenciana (Prometeo II-2014-053).

†de Lucio: Universidad Nebrija. Calle de Santa Cruz de Marcenado, 27, 28015, Madrid (Spain). Email:jlucio@nebrija.es.

‡Mínguez: Universidad Nebrija. Calle de Santa Cruz de Marcenado, 27, 28015, Madrid (Spain). Email:rminguez@nebrija.es.

§Minondo: Corresponding author. Deusto Business School, University of Deusto, Camino de Mundaiz 50, 20012 Donostia - San Sebastián (Spain). Research affiliate of Instituto Complutense de Estudios Internacionales. Email: aminondo@deusto.es

¶Requena: Department of Economic Structure, University of Valencia, Avda. dels Tarongers s/n, 46022 Valencia (Spain). Email: francisco.requena@uv.es

1 Introduction

Exports are dominated by few firms. For example, the top 1,000 firms in Spain explained two-thirds of exports in 2015. As stated by Gabaix (2011), these large firms, also known as superstars, are the incompressible grains of economic activity. If few firms account for most of exports, they should play a crucial role in shaping countries export specialization and dynamics. In this paper, we use Spanish firm-level export data to document this pivotal role.

We show first that exports are more concentrated by firm than sales or employment. We also observe that this concentration has not changed substantially over the period 1997-2015. Moreover, we document heterogeneity in export concentration across products. Second, export specialization is dominated by superstars. If we removed the top 10 firms, Spain would lose revealed comparative advantage in 56% of HS 2-digit chapters, accounting for 26% of total exports. This result suggests that in many industries, revealed comparative advantage is explained by highly successful firms, rather than fundamental, or country-level, variables. Finally, we show that superstars can explain around one-third of the variation in aggregate exports.

This paper is related to the recent empirical literature analyzing the granularity of exports. In particular, it is related to Freund and Pierola (2015) that analyze export concentration in 32 developing countries, and investigate whether countries' trade patterns are defined by superstars.[1] Our research adds to the literature analyzing the weight and role of top exporters in Spain, a major exporter in the world. We also add to the literature showing that export concentration is stable throughout time and occurs even at finely defined export operations. We also document heterogeneity in export granularity across products, specially in more aggregated classifications. Finally, we measure the idiosyncratic contribution of top firms to the variation of exports.

In a broader sense, our research is also related to the literature that has introduced granularity into trade models (Eaton et al., 2012; Bernard et al., 2016), and the literature that has applied these models to estimate the role of fundamental and granular forces in shaping comparative advantage (Gaubert and Itskhoki, 2016). Our research provides empirical support for this class of models. Finally, our paper is also related to the literature that has analyzed how granularity might affect important economic phenomena, such as aggregate volatility (Gabaix, 2011; Di Giovanni and Levchenko, 2012; Di Giovanni et al., 2014), welfare (Di Giovanni and Levchenko, 2013), or trade balance (Canals et al., 2007). We add to this literature investigating the idiosyncratic contribution of superstars to the variation of exports.

[1] Other papers, such as Canals et al. (2007), Mayer and Ottaviano (2008), Bernard et al. (2009) and Marin et al. (2015) also document export concentration by firms.

The rest of the paper is split into four different sections. Section 2 describes the concentration of exports by firm. Section 3 analyzes whether superstars define the specialization of Spanish exports. Section 4 investigates the contribution of top firms to the fluctuations in Spanish exports. Section 5 presents the main conclusions of the paper.

2 The concentration of exports

2.1 Database

To calculate the concentration of Spanish exports, we use the universe of export transactions database, which is elaborated by the Customs and Excise Department of the Spanish Tax Agency. For each transaction, we know the firm's identification code, the product at the 8-digit Combined Nomenclature (CN) classification[2], the destination of the export transaction, the free-on-board (FOB) value in euros of the transaction, and the exported quantity (in weight metric and/or units).[3] Every year new CN8 codes are added to the classification and some CN codes are dropped. Since some sections of the paper combine data for different years, we use the Van Beveren et al. (2012) algorithm to create a consistent product classification over the period of analysis.[4] With this database, we analyze, first, how export concentration varies by the number of top firms. Second, we study how sensitive concentration is to more detailed export operations and different product classifications.

2.2 Concentration of exports by firm

We begin analyzing the concentration of exports by firm. Table 1 presents the concentration of exports for the top firm, top 5, top 50, top 200 and top 1,000 firms over the period 1997-2015. We use absolute numbers to identify top firms, rather than percentages (e.g. top 1% of exporters), because the number of exporters increases with GDP (Fernandes et al., 2015).

In 2015, the top Spanish firm accounted for 3% of all exports. This firm exported 3,725 times more than the average exporter, and almost 2 million times more than the median exporter. If we remove the very small export operations, ($< 6,000$ euros)[5], the top Spanish firm exported 939 times more than the average exporter, and 10,354 times more than the median exporter.

[2] An example of an 8-digit product is CN 87120030 Bicycles with ball bearings.

[3] Due to their special geographical situation and fiscal status, we remove the Spanish firms registered in the two autonomous cities located in Africa (Ceuta and Melilla).

[4] We have a consistent classification for 7,650 product categories over the period 1997-2015.

[5] Up to this value European Union (EU) exporters do not have to certify that the product meets EU's rules of origin http://madb.europa.eu/madb/rulesoforigin_preferential.htm

The top 5 firms represented 10% of all exports, the top 50 almost one-third of all exports, the top 200 almost half of all exports, and the top 1000 two-thirds of all exports. These figures point out that the concentration of exports by firm is very high. If we removed the top 1,000 firms from the Spanish economy, which only represented 0.7% of all exporting firms in 2015, 67% of all exports would vanish. Table 1 also shows that over the period 1997-2015 there has been a moderate reduction in the concentration of exports in the top 5 firms. In contrast, the concentration of exports is similar for the top 1 and top 50 firms, and rises slightly for the top 200 and top 1,000 firms.

Export concentration by firm is higher in merchandises than in services. Minondo (2013) shows that on average, over the period 2001-2007, the top 1% of firms account for 49% of service exports, the top 5% of firms for 71% of service exports, and the top 10% of firms for 80% of service exports. For merchandises, the figures for the same period are 63%, 85% and 93% for the top 1%, 5% and 10% respectively.

The concentration of exports by firm is also higher than the concentration of total sales or employment by firm. We use the Bureau Van Dick SABI database, which provides detailed financial and accounting records of Spanish firms that deposited their accounts in the Business Register. It is important to point out that small firms are under-represented in SABI. Hence, our calculations provide upper-bound estimates for the concentration of output and employment.

Table 2 presents the concentration of sales and employment by firm in Spain in 2013. To facilitate the comparison, we include export concentration by firm in that year. Concentration in exports is much higher than in sales (which includes sales in domestic and foreign markets) and employment. For example, whereas the top firm exported 3% of all Spanish exports, the top firm in sales and employment only accounted for 1% of those variables. Regarding the top 1,000 firms, they made up more than two-thirds of exports, whereas they accounted for less than half of sales, and less than a third of employment.[6] The higher concentration in export than in sales and employment might be explained by the magnification effect of trade. As argued by Bernard et al. (2016), small differences in productivity across firms lead to magnified differences in exports due to the interdependence across margins. For example, slightly more productive firms can source inputs from more markets, which grants them a competitive edge to raise exports in their leading market, which generates the resources to overcome fixed costs to enter new destinations and products, or to shift production to more competitive locations.

Concentration of exports by firm is not a particular characteristic of the Spanish economy. Freund and Pierola (2015) show that among 32 developing countries, and excluding oil exports, the top firm accounts, on average, for 14% of exports, and the top 5 firms for 30% of all exports over the period 2006-2008. These concentrations are larger

[6] Bernard et al. (2009) also document a higher granularity in exports than in employment for the US.

Table 1: Share of superstars in Spanish exports, 1997-2015

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Top 1	3	4	4	4	3	3	3	3	3	3	3	3	3	2	2	3	3	2	3
Top 5	13	14	13	12	12	11	12	13	11	11	11	10	11	10	9	10	10	10	10
Top50	32	31	31	31	29	28	29	30	31	32	32	30	30	31	32	32	33	33	32
Top 200	45	44	43	44	42	41	42	44	45	46	46	45	45	46	48	48	49	48	48
Top 1,000	65	64	64	64	63	62	64	65	65	66	66	66	66	67	68	68	68	68	67

Source: Own elaboration using the Customs database.

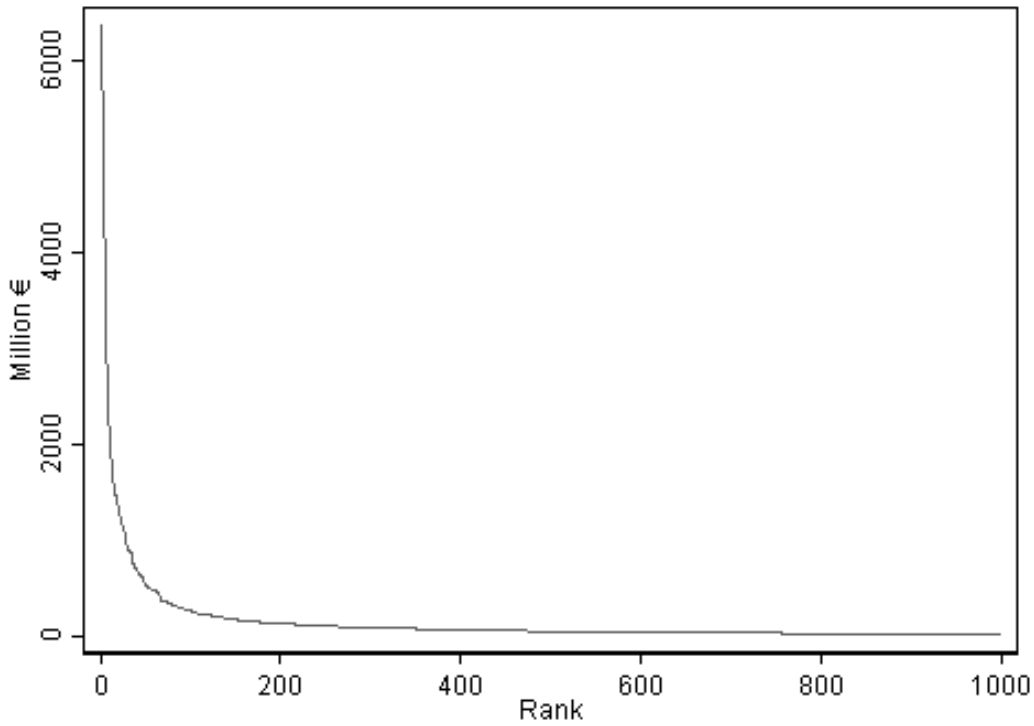
Table 2: Concentration in exports vs. output and employment, 2013 (% of total)

	Top 1	Top 5	Top 50	Top 200	Top 1,000
Exports	3	10	33	49	68
Sales	1	6	20	31	46
Employment	1	3	10	18	29

Source: Authors calculations from SABI and Customs database.

than the ones we have found for Spain. The concentration of exports in Spain is more similar to France, where the top 1% of exporters, roughly equivalent to the top 1,000 firms in Spain in 2015, accounted for 68% of all exports (Mayer and Ottaviano, 2008).

Figure 1: The distribution of exports in the top 1,000 firms, 2015



Source: Own elaboration using the Customs database.

Figure 1 presents the distribution of exports in the top 1,000 firms in 2015. We rank the top exporting firm as first in the ranking. The figure shows that exports drop dramatically as we descend in the ranking. Using the methodology proposed by Gabaix and Ibragimov (2011), we test whether a Pareto distribution fits the distribution of exports in the top 1,000 Spanish firms. The regression concludes that 99.5% of the differences in the (log) rank of firms is explained by the (log) exports. This result confirms that a Pareto distribution fits very well the distribution of exports in the top 1,000 exporters.[7]

[7]The regression yields a Pareto parameter equal to 1.1

To finish this first set of analyses, we investigate the differences between the top 1,000 firms and the rest of exporters. As shown in Table 3, superstars export to more destinations and more products than ordinary firms.[8] In particular, the median superstar exports to 30 different destinations, whereas the median ordinary firm only exports to one destination. However, there are superstars that only export to one destination; and there are a few ordinary firms that export to many destinations. Among superstars, the maximum number of destinations is 141.

Differences between superstars and ordinary exporters are also salient regarding products. Among superstars, the median firm exported 21 products in 2015, whereas the median ordinary firm only exported one. As before, we observe large differences across firms within each group. Among superstars there are firms that export up to 1,946 products and firms that only export one product. Among ordinary firms, while most firms only export one product, there is a firm that exports up to 1,420 different products.

Table 3: Superstar exporters vs. ordinary exporters, 2015

	Firms	Median	Mean	Std. dev.	Minimum	Maximum
Destinations						
Superstars	1,000	30	38	29	1	141
Ordinary	146,118	1	3	7	1	116
Products						
Superstars	1,000	21	57	129	1	1,946
Ordinary	146,118	1	5	15	1	1,420

Std. dev. =Standard deviation. Source: Own elaboration using the Customs database.

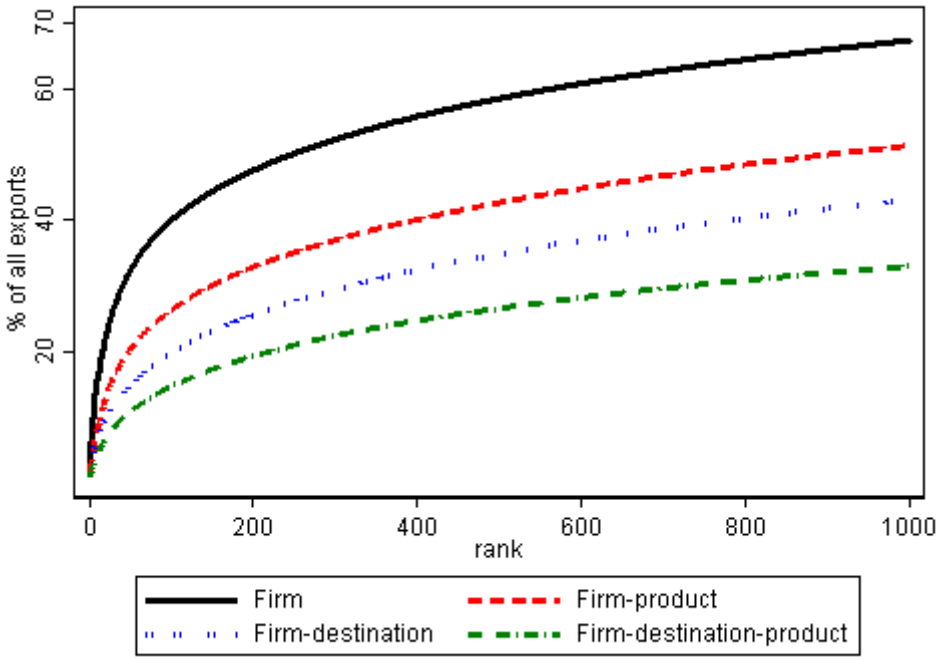
2.3 Zooming in on concentration

The customs database provides export value by firm, product and destination. This allow us to zoom in on the concentration of exports. In particular, we want to analyze how the concentration of exports differs when we zoom in from exports by firm, to exports by firm-destination, exports by firm-product and exports by firm-destination-product. Figure 2 presents the results of the zoom-in exercises. As benchmark, we draw the accumulated percentage of exports up to the top 1,000 firms (solid black line). This line reproduces the figures presented in Table 1 for 2015.

The top firm-product dyad represents 1.1% of all Spanish exports, the top firm-destination dyad 1.0% of all exports, and the top firm-destination-product triad 0.7% of all exports. This compares to the 3% of all Spanish exports accounted by the top Spanish firm. Although the zoom-in process reduces the concentration of exports, it is

[8]Every year new CN8 codes are added to the classification and some CN codes are dropped from the classification. We use the Van Beveren et al. (2012) algorithm to create a consistent product classification.

Figure 2: Concentration of exports for different levels of aggregation, 2015



Source: Own elaboration using the Customs database.

remarkable that the top firm-destination-product triad still represents almost 1% of all Spanish exports.[9]

The difference in concentration between firms and our zoom-in categories widens as we move to the right in the horizontal axis. However, it is still remarkable the concentration of the zoom-in categories for the top 1,000 dyads or triads. For example, the top 1,000 firm-product dyads represent more than 51% of Spanish exports, compared to the 67% for the top 1,000 firms. For the firm-destination dyad the figure is 43% and for the firm-destination-product triad 33%. These results point out that exports are not only concentrated by firms, but also by destinations and products within each firm.

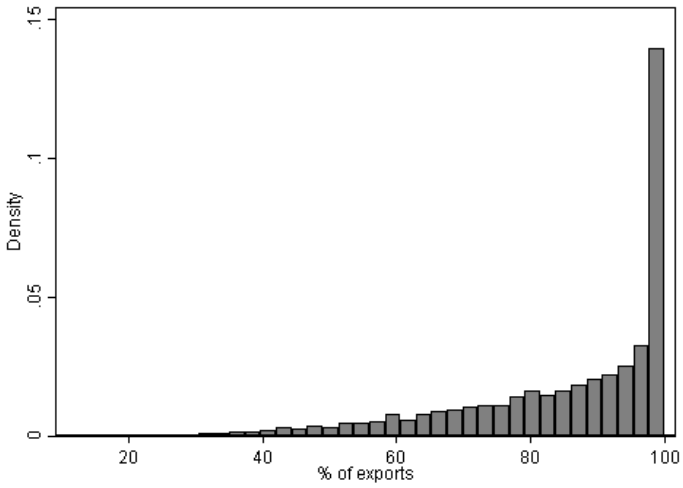
2.4 Concentration across product classifications

In this subsection we analyze whether the concentration of exports by firm is homogeneous across products. We perform the analysis for three different product classifications. First, we use the 8-digit CN classification, consistent across time, which defines 7,650 products. Second, we aggregate exports at the HS 2-digit level, which defines 96 chapters. Finally, we aggregate exports at the section level, which defines 21 categories. The latter two

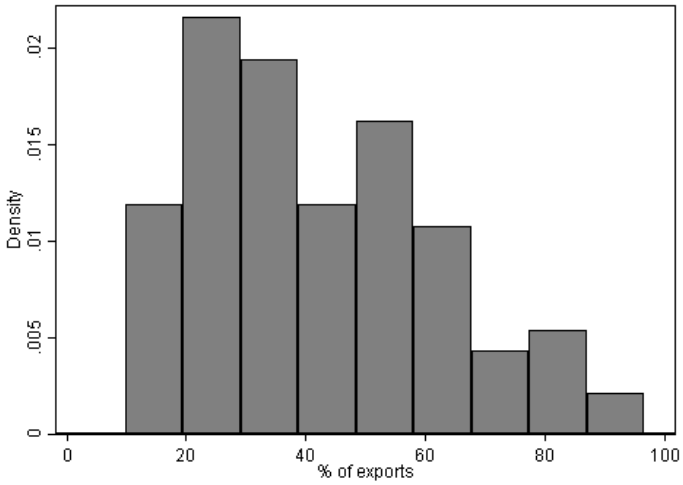
[9]It is interesting to note that the top firm-destination-product triad in 2015 did not belong to the top exporter.

Figure 3: Concentration of exports in the top 5 firms across products, 2015

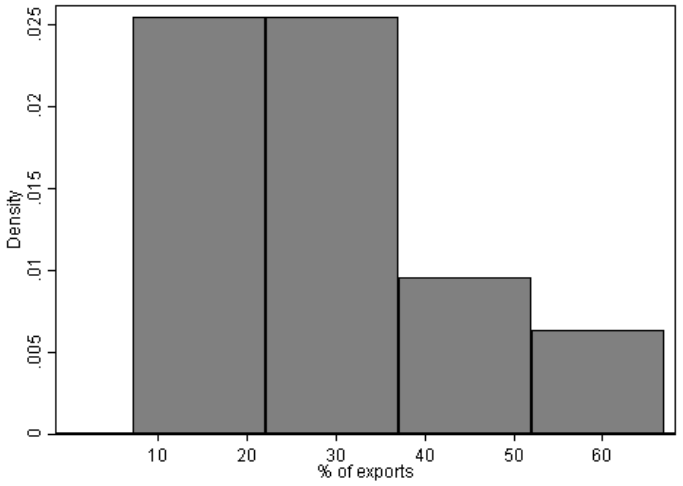
(a) CN 8-digit classification (7,560 products)



(b) Export chapters (96 chapters)



(c) Export sections (21 sections)



Source: Own elaboration using the Customs database.

classifications are used in the next section, when we analyze the influence of top firms in Spanish export specialization. To perform these analyses we use data for the year 2015, and calculate the share of the top five firms in each product's total exports.

Figure 3 presents the histograms for the three different product disaggregations. In the CN classification, panel (a), the top five firms account for almost 100% of exports in the majority of products. In particular, in the median product, the share of the top five firms in total exports is 91%. In panel (b), we define products at the chapter level. Since products are much more aggregated in this classification, the share of the top five firms in total exports is reduced significantly. In the median chapter, the share of the top five firms is 38%. We also observe a larger heterogeneity across chapters. For example, the share of chapters where the top five firms account between 10% and 20% of exports is similar to the share of chapters where the top five firms account between 40% and 50% of exports. The concentration is even lower for sections (panel (c)), where the top five firms account for 25% of exports in the median section.

3 Do superstars determine the specialization of Spanish exports?

Following Freund and Pierola (2015), we analyze how Spanish revealed comparative advantage (RCA) would alter if we removed superstar firms. To do so, we calculate the Balassa index, which is defined as follows:

$$RCA_k = \frac{x_{esp,k}/X_{esp}}{x_{world,k}/X_{world}} \tag{1}$$

where $x_{esp,k}$ and $x_{world,k}$ are Spanish and World exports of product k respectively; and X_{esp} and X_{world} are total Spanish and World exports respectively.

We calculate the Balassa index for each of the 96 HS 2-digit chapters. We select this classification to ensure a large enough number of firms per category.[10] To attenuate year-specific effects, we pool data for 2010, 2011 and 2012. First, we calculate the Balassa index with all Spanish firms. Then, for each chapter, we re-calculate the index excluding the chapter-specific superstar firms. We use the BACI database to calculate world export aggregates (Gaulier and Zignago, 2010).[11]

Table 4 shows the results of the calculations. First, we present the RCA index for all chapters. The chapter with the highest RCA index is cork and articles of cork (HS2 code 45). Spain also has high RCA indexes in the chapters related with food and food industry, such as edible vegetables, edible fruit, vegetable and fruit preparations, and beverages.

[10] All chapters have, at least, 48 firms exporting above 18,000 euros in the 2010-2012 period.
[11] We exclude chapters 98 and 99, since they are not included in the BACI database.

Spain also shows a high RCA in the ceramic industry, in railways and automobiles, in pharmaceutical products, in apparel and footwear, and iron and steel products. Note that automobiles is the chapter that accounts for the largest share in exports (%16.84) and pharmaceuticals is the fifth largest (%4.54).

The RCA All column identifies the chapters with a clear RCA (≥ 1.1) and with a border RCA ($0.9 < RCA < 1.1$) when all firms are included. The chapters left in blank have a RCA ≤ 0.9. There are 45 chapters, over a total of 96, where Spain has a clear RCA; 8 chapters have a border RCA and 43 chapters do not have RCA. The next column, RCA no Top 1, identifies whether the chapter would lose its RCA if the top 1 firm was removed from the market. The number of chapters with a clear RCA drops to 31. However, 11 chapters move to a border RCA, and only 3 chapters clearly lose their RCA.[12] Hence, only 7% of chapters (3 over 45) clearly lose their RCA when we remove the top 1 firm. However, the influence of top firms on RCA increases as we widen the set of superstar firms. When we exclude the top 5 firms, the number of chapters where Spain commands a clear RCA drops from 45 to 20, with seven of these industries moving to the border RCA situation. When we enhance the set of superstars to the top 10 firms, the number of industries with RCA drops from 45 to 16, with four of these industries moving to the border situation. Half of the chapters that keep their RCA belong to the food industry. Hence, the top 5 firms determine the RCA in 40% of chapters (18 chapters clearly losing their RCA over 45 chapters), and the top 10 firms in 56% of chapters (25 chapters losing their RCA over 45 chapters). The chapters in which Spain loses RCA when the top 10 firms are removed account for 26% of all Spanish exports, and 40% of exports of the chapters in which Spain has RCA.[13]

These percentages suggest that superstars have a very large influence on export specialization. However, we should qualify this statement, since we do not know how the remaining firms would behave if superstars disappeared. Using a different methodology, Gaubert and Itskhoki (2016) analyze the contribution of granular (superstars-based) comparative advantage and fundamental (country-based) comparative advantage to differences in the share of exports across French industries. They find that granular comparative advantage explains one-third of the differences, and fundamental comparative advantage explains two-thirds of the differences.

[12] These are rubber; pulp of wood; and articles of apparel and clothing.

[13] As robustness check, we also transform the RCA variable into a revealed symmetric comparative advantage (RSCA) variable (Laursen, 2015). The RSCA is defined as (RCA-1)/(RCA+1) and has a $[-1, 1]$ range. Using a RSCA ≥ 0.1 threshold to determine a clear RCA, we obtain very similar results.

Table 4: The contribution of superstars to RCA. Balassa Index

HS2 code	Description	% of exports	RCA index	RCA All	RCA No Top1	RCA no Top 5	RCA no Top 10
1	Live animals	0.19	1.68	Yes	Yes	Border	
2	Meat and edible meat	1.67	2.56	Yes	Yes	Yes	Yes
3	Fish	1.07	1.93	Yes	Yes	Yes	Yes
4	Dairy products	0.49	1.02	Border	Border		
5	Products of animal origin	0.09	1.74	Yes	Yes	Yes	
6	Live tree and other plants	0.12	1.04	Border	Border		
7	Edible vegetables	2.05	5.85	Yes	Yes	Yes	Yes
8	Edible fruit	2.88	5.36	Yes	Yes	Yes	Yes
9	Coffee and tea	0.16	0.59				
10	Cereals	0.19	0.26				
11	Milling products	0.10	0.98	Border			
12	Oil seed	0.20	0.43				
13	Lac, gums and resins	0.10	2.31	Yes	Yes	Border	
14	Vegetable products	0.00	0.41				
15	Animal and vegetable fats and oils	1.37	2.12	Yes	Yes	Yes	Yes
16	Preparations of meat and fish	0.51	2.00	Yes	Yes	Yes	Yes
17	Sugars and sugar confectionery	0.24	0.76				
18	Cocoa and cocoa preparations	0.25	0.89				
19	Preparations of cereal	0.47	1.41	Yes	Yes		
20	Vegetable, fruit, and food preparations	1.07	3.44	Yes	Yes	Yes	Yes
21	Miscellaneous edible preparations	0.57	1.73	Yes	Yes	Border	
22	Beverages	1.57	2.60	Yes	Yes	Yes	Yes
23	Residues and waste from the food industry	0.34	0.89				
24	Tobacco	0.14	0.49				
25	Salt, sulphur, earth, stone, plaster, lime and cement	0.51	1.76	Yes	Yes	Yes	Yes
26	Ores and slag	0.63	0.44				
27	Mineral fuels	6.64	0.27				
28	Inorganic chemicals	0.45	0.56				
29	Organic chemicals	1.82	0.65				
30	Pharmaceutical products	4.54	1.69	Yes	Yes		
31	Fertilisers	0.34	0.79				
32	Tanning/dyeing extracts	0.97	2.19	Yes	Yes	Yes	Yes
33	Essential oils and resinoids; perfumery and cosmetics	1.26	2.19	Yes	Yes	Yes	Yes
34	Soap	0.48	1.50	Yes	Yes	Yes	Border

Continued on next page

Table 4 – *Continued from previous page*

HS2 code	Description	% of exports	RCA index	RCA All	RCA no Top 1	RCA no Top 5	RCA no Top 10
35	Albuminoidal substitutes, starches, glues and enzymes	0.06	0.40				
36	Explosives	0.01	0.26				
37	Photographic or cinematographic goods	0.06	0.59				
38	Miscellaneous chemical products	1.09	1.10	Border	Border		
39	Plastics	3.95	1.22	Yes	Yes		
40	Rubber	1.71	1.36	Yes			
41	Raw hides and skins	0.32	1.87	Yes	Yes	Yes	Yes
42	Articles of leather	0.31	0.93	Border			
43	Furskins and artificial fur	0.06	1.19	Yes	Border		
44	Wood and articles of wood	0.52	0.77				
45	Cork and articles of cork	0.10	10.06	Yes	Yes	Yes	Yes
46	Manufactures of straw	0.00	0.36				
47	Pulp of wood and fibrous cellulosic material	0.41	1.42	Yes			
48	Paper and paperboard	1.64	1.51	Yes	Yes	Border	
49	Printed books, newspapers and pictures	0.35	1.31	Yes	Yes	Border	Border
50	Silk	0.01	0.59				
51	Wool	0.05	0.55				
52	Cotton	0.31	0.85				
53	Other vegetable textile fibres	0.01	0.64				
54	Man-made filaments	0.25	1.00	Border			
55	Manmade staple fibres	0.17	0.75				
56	Wadding, felt and nonwoven; yarns; twine, cordage	0.17	1.36	Yes	Border		
57	Carpets and other textile floor coverings	0.03	0.39				
58	Special woven fabrics	0.08	1.18	Yes	Border		
59	Impregnated, coated, cover/laminated textile fabric	0.15	1.17	Yes	Border		
60	Knitted or crocheted fabrics	0.12	0.82				
61	Articles of apparel and clothing, knitted or crocheted	1.38	1.20	Yes			
62	Articles of apparel and clothing, not knit or crochet	1.86	1.61	Yes	Border		
63	Other made up textile articles	0.19	0.61				
64	Footwear	1.04	1.58	Yes	Yes	Yes	Border
65	Headgear	0.03	0.61				
66	Umbrellas	0.01	0.57				
67	Feathers, artificial flowers, and articles of human hair	0.01	0.22				
68	Articles of stone, plaster, and cement	0.58	2.43	Yes	Yes	Yes	Yes
69	Ceramic products	1.12	4.54	Yes	Yes	Yes	Yes

Continued on next page

Table 4 – *Continued from previous page*

HS2 code	Description	% of exports	RCA index	RCA All	RCA no Top 1	RCA no Top 5	RCA no Top 10
70	Glass and glassware	0.50	1.23	Yes	Border		
71	Pearls, precious stones, and metals	1.54	0.53				
72	Iron and steel	3.65	1.41	Yes	Yes	Border	
73	Articles of iron or steel	2.65	1.56	Yes	Yes	Yes	Yes
74	Copper	1.20	1.06	Border			
75	Nickel	0.01	0.07				
76	Aluminium	1.15	1.23	Yes	Border		
78	Lead	0.02	0.37				
79	Zinc	0.35	4.00	Yes	Yes		
80	Tin	0.02	0.21				
81	Other base metals	0.02	0.16				
82	Tools, implements and cutlery	0.31	0.93	Border			
83	Miscellaneous articles of base metal	0.48	1.40	Yes	Yes	Border	
84	Nuclear reactors, boilers, machinery	7.90	0.67				
85	Electrical and electronic equipment	6.26	0.53				
86	Railways	0.51	2.60	Yes	Border		
87	Vehicles other than railways	16.84	2.31	Yes	Yes	Yes	Border
88	Aircrafts	1.56	1.35	Yes	Border		
89	Ships	0.60	0.73				
90	Optical, photo, technical, medical, etc apparatus	0.92	0.30				
91	Clocks and watches	0.13	0.47				
92	Musical instruments	0.01	0.37				
93	Arms and ammunition	0.06	1.29	Yes	Border		
94	Furniture; bedding, mattress, material support and cushion	0.99	0.88				
95	Toys, games and sports requisites	0.36	0.66				
96	Miscellaneous manufactured articles	0.20	1.17	Yes	Border		
97	Works of art, collectors' pieces and antiques	0.04	0.33				
	Industries with RCA$\geq$ 1.1			45	31	20	16
	Industries with $0.9 < RCA < 1.1$			8	11	7	4
	Industries with RCA $\leq$ 0.9			43	54	69	76

Note: RCA index is calculated when all firms are included. "Yes" denotes that the HS2 chapter has a RCA $\geq$ 1.1. "Border" denotes that the HS2 chapter has a RCA between 0.9 and 1.1. If blank, the HS2 chapter has a RCA $\leq$ 0.9. RCA no Top 1, RCA no Top 5, RCA no Top 10 are calculated removing the top 1, top 5 and top 10 firms from the chapter respectively. Source: Own elaboration using the Customs database.

The contribution of granular comparative advantage is similar to the contribution of the top 5 firms to RCA in Spain.[14]

As a robustness exercise, following Costinot et al. (2012), we calculate a regression-based RCA index, consistent with the Ricardian model of trade. Specifically, we estimate the following regression:

$$\ln x_{ijk} = \delta_{ij} + \delta_{ik} + \delta_{jk} + \epsilon_{ijk} \tag{2}$$

where δ_{ij}, δ_{ik} and δ_{jk} are exporter-importer, exporter-product and importer-product fixed effects.

As explained by Costinot et al. (2012), the estimated exporter-product fixed effects are theoretically consistent estimates of the productivity of exporter i in product k. It is important to note that, due to the properties of the linear regression, all the estimated exporter-product productivities are defined relative to a numeraire good and a numeraire exporter. Hence, exporter-product fixed effects capture the productivity of exporter i in product k relative to the productivity of a numeraire exporter i' in a numeraire product k'.[15] It is important to note that μ_{ik} are estimates of revealed measures of relative productivity, because it is only possible to estimate the relative productivity of exporter i in product k if exporter i exports product k.[16]

The limitation of the regression procedure is that computational capacity runs into difficulties if the number of fixed effects is high. To reduce this number, we group HS 2-digit chapters into the 21 HS sections. Next, we restrict the sample to the top 40 exporter countries in the world, which account for 77% of world exports. Once we have estimated country-industry fixed effects, following Freund and Pierola (2015), we estimate the regression based RCA ($RCAr$) using:

$$RCAr_k = \frac{X_{esp,k}/mean(X_{esp,k})}{mean(X_{j,k}/mean(X_{j,k}))} \tag{3}$$

where $X_{esp,k}$ is the exponent of the fixed effect of Spain in industry k. This variable captures Spanish exports in industry k relative to the numeraire country and product; $mean(X_{esp,k})$ is the average relative export across industries in Spain. In the denominator, for each country, we divide relative exports in industry k by the average relative exports

[14] Gaubert and Itskhoki (2016) methodology relies on the distribution of sales per firm in the domestic and foreign markets. We cannot apply this methodology because we do not have information about domestic sales per firm.

[15] Following Freund and Pierola (2015) we select a section with a low granularity level, and where Spain does not have RCA, as the numeraire industry: machinery. We select USA as the numeraire country.

[16] We could use the exporter-importer fixed effects to calculate $(\mu_{ik} - \mu_{ik'}) - (\mu_{i'k} - \mu_{i'k'})$. This expression captures the Ricardian comparative advantage, since it compares the productivity of exporter i in product k relative to product k' with the same ratio in exporter i'.

in the rest of industries and, then, we take the average of this ratio over all countries.

Table 5 presents the results of the analysis with the regression-based RCA index (RCAr). For comparison, we also present the results obtained with a Balassa-based RCA index. As before, we consider that Spain has a clear RCA in industry k if RCA is ≥ 1.1, a border RCA if $0.9 < RCA < 1.1$, and clearly no RCA if RCA≤ 0.9. According to the regression-based methodology, Spain has a clear comparative advantage in 10 sections over a total of 20.[17] In line with Table 4, Spain has a RCA in vegetables, fats and oils, foodstuff and beverages, chemicals, leather, textile and garments, footwear, stone and cement (which includes ceramics), transport equipment and arms. When we remove the top 1 firm, the number of sections with a clear RCA drops to 7. When we remove the top 5 firms the number of sections with a clear RCA drops to 6; and when we remove the top 10 firms the number of sections with a clear RCA drops to 3. The sections that keep their RCA after this last removal are foodstuff and beverages, footwear and stone and cement. When we remove the top 10 firms, four sections shift from a clear RCA to a border situation: vegetables, fats and oils, leather and transport industry. And three sections clearly lose their RCA: chemicals, textiles and garments, and arms. According to these figures, when we remove the top 10 firms, Spain loses clearly the RCA in 30% (3 over 10) of sections. These sections represent 16% of all Spanish exports, and 29% of exports in which Spain has RCA.

The last four columns present the results when RCA is computed with the Balassa index. When all firms are included, Spain has a RCA in 13 sections over 21. When we remove the top 10 firms, Spain keeps a clear RCA in 6 of these 13 sections; 2 of them move to the border range, and 5 clearly lose their RCA. According to the Balassa index, Spain would lose clearly the RCA in 38% of sections (5 over 13). This figure is higher than the one obtained with the regression-based calculations. However, it is lower than the percentage we obtain when the analysis was performed at the chapter level (Table 4). Since there are more exporters per industry at the section than at the chapter level, the removal of the top 10 firms has a lower impact in the former than in the later. The sections losing the RCA when the top 10 firms are removed represent 25% of all Spanish exports, and one-third of exports in which Spain has RCA.

4 Granularity and the dynamics of exports

In this section we analyze whether superstars shape the dynamics of Spanish exports. To do so, we follow the methodology proposed by Gabaix (2011).[18] The change in exports

[17] Note that we cannot calculate the RCA in machinery, since it is the numeraire section.

[18] Canals et al. (2007) and Di Giovanni et al. (2014) also develop methodologies to measure the impact of firm-specific shocks to variations in exports.

Table 5: Superstars and RCA in Spanish industries

Section code	Name	% exports	(a) Regression-based RCA				(b) Balassa RCA			
			RCA All	RCA no Top 1	RCA no Top 5	RCA no Top 10	RCA All	RCA no Top 1	RCA no Top 5	RCA no Top 10
I	Animals	3.16	Border				Yes	Yes	Yes	Yes
II	Vegetables	6.81	Yes	Yes	Border	Border	Yes	Yes	Yes	Yes
III	Fats and oils	1.42	Yes	Yes	Yes	Border	Yes	Yes	Yes	Yes
IV	Foodstuff and beverages	5.09	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
V	Minerals	4.81								
VI	Chemicals	11.95	Yes	Border			Yes	Border	Border	
VII	Plastic and rubber	5.76	Border				Yes	Border		
VIII	Leather	0.73	Yes	Yes	Yes	Border	Yes	Yes	Border	
IX	Wood	0.48								
X	Paper	2.24	Border	Border			Yes	Yes		
XI	Textiles and garments	4.20	Yes	Border			Yes			
XII	Footwear	1.13	Yes	Yes	Yes	Yes	Yes	Yes	Border	Border
XIII	Stone and cement	2.00	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
XIV	Precious metals	2.04								
XV	Base metals	9.50	Border	Border			Yes	Yes	Border	Border
XVI	Machinery	13.94	NA	NA	NA	NA				
XVII	Transport equipment	22.31	Yes	Yes	Yes	Border	Yes	Yes	Yes	Yes
XVIII	Optical instruments	0.98								
XIX	Arms	0.07	Yes				Yes	Border		
XX	Other manufactures	1.33								
XXI	Works of art	0.05								
Industries with RCA$\geq$ 1.1			10	7	6	3	13	10	6	6
Industries with $0.9 < RCA < 1.1$			4	4	1	4	0	2	4	2
Industries with RCA ≤ 0.9			6	9	13	13	8	9	11	13

Note: NA (Not Available). "Yes" denotes that the HS2 section has a RCA ≥ 1.1. "Border" denotes that the HS2 section has a RCA between 0.9 and 1.1. If blank, the HS2 section has a RCA ≤ 0.9. RCA no Top 1, RCA no Top 5, RCA no Top 10 are calculated removing the top 1, top 5 and top 10 firms from the section respectively. Source: Own elaboration using the Customs database.

of firm f selling product k in market d can be separated into three components: a firm-specific, or idiosyncratic, component; a product component; and a destination component. It is the first component which captures the contribution of superstars to the variation in aggregate exports. To separate the idiosyncratic component, first, we calculate the growth in exports for a specific firm, product and destination:

$$g_{f,k,d,t} = log(x_{f,k,d,t}) - log(x_{f,k,d,t-1}) \tag{4}$$

where $x_{f,k,d,t}$ are exports by firm f of product k in destination d at time t. Note that (4) can only compute export transactions at the firm, product and destination level that happen in year $t-1$ and year t. Hence, it only captures the intensive margin of trade. This is a reasonable comprise, providing that the intensive margin explains around 77% of the year-by-year variation in exports in Spain (de Lucio et al., 2011).

Then, we calculate the average of the growth rates of all firms in product k and in market d:

$$\bar{g}_{k,d,t} = \frac{\sum_f g_{f,k,d,t}}{N_{k,d,t-1}} \tag{5}$$

where $N_{k,d,t-1}$ is the total number of exporters of k in destination d at year $t-1$. Note that $g_{k,d,t}$ aggregates the product-level and destination-level growth components.

Then, we identify the idiosyncratic component, $ic_{f,k,d,t}$, as the difference between the firm growth rate and the average growth rate:

$$ic_{f,k,d,t} = g_{f,k,d,t} - \bar{g}_{k,d,t} \tag{6}$$

Finally, we calculate the granular residual as the sum of the idiosyncratic components of top firms, weighted by the share of each superstar in total exports in product k and market d at time $t-1$:

$$r_{k,d,t} = \sum_{f \in S_{k,d,t-1}} ic_{f,k,d,t} s_{f,k,d,t-1} \tag{7}$$

where $s_{f,k,d,t-1}$ is the share of firm f in total exports of product k in destination d at year $t-1$, and $S_{k,d,t-1}$ is the set of superstar firms in product k and market d at time $t-1$.

To measure the contribution of superstar firms to the changes in aggregate exports, we regress the change in aggregate exports by product and destination on the granular residual. Following Gabaix (2011), in addition to the contemporary granular residual, we add granular residual 1-year and 2-years lagged values to control for dynamic inter-

Table 6: Contribution of superstars to variations in Spanish exports, 1997-2015

	(1)	(2)	(3)	(4)
Granular residual	0.633***	0.659***	0.577***	0.608***
	(0.007)	(0.008)	(0.006)	(0.006)
Granular residual 1 year lag		-0.103***		-0.125***
		(0.006)		(0.005)
Granular residual 2 years lag		-0.064***		-0.088***
		(0.006)		(0.005)
Constant	0.063***	0.041***	0.111***	0.068***
	(0.001)	(0.001)	(0.001)	(0.002)
Superstars	Top 1 firm	Top 1 firm	Top 5 firms	Top 5 firms
Observations	62589	50353	62589	50353
Adjusted R square	0.287	0.309	0.328	0.369

Note: Robust standard errors in parentheses. ***, **, * statistically significant at 1%. 5% and 10% respectively. Data source: Customs database.

dependencies. The regression equation is the defined as follows:

$$g_{k,d,t} = \alpha + \beta_1 r_{k,d,t} + \beta_2 r_{k,d,t-1} + \beta_3 r_{k,d,t-2} + \epsilon_{k,d,t} \tag{8}$$

where α is a constant and $r_{k,d,t-1}$ and $r_{k,d,t-2}$ are granular residual 1 and 2 year lags respectively. We define products at the HS 2-digit disaggregation and restrict the sample to those product and destination combinations that have at least 10 exporters operating in year $t-1$ and year t.[19] We analyze the contribution of the top 1 firm and the contribution of the top 5 firms to the variation of exports.

Table 6 presents the results of the regression analyses. The granular residual coefficients for the top 1 firm and the top 5 firms are positive and highly statistically significant. These coefficients point out that superstars influence the variation in exports. In particular, according to the adjusted R-square statistic in column (1), the top 1 firm explains 29% of the variation in exports; according to column (4), the top 5 firms explain 33% of the variation. The coefficient for the contemporary granular residual remains positive and statistically significant once we control for lagged values.

For output growth, Gabaix (2011) finds that the top 100 firms in the US explain around one-third of the variations over the period 1951-2008. Wagner (2012) finds that the top 10 firms explain between 36% and 45% of the fluctuation of manufacturing sales

[19]To minimize the impact of outliers, we winsorize growth rates to the 5% and 95% percentiles.

in Germany between 2007 and 2008. These results are in the range of the contribution we have estimated for Spain.[20]

5 Conclusion

Heterogeneity among firms is outstanding in exports. In Spain the top 200 firms accounted for half, and the top 1,000 firms for two-thirds, of Spanish merchandise exports in 2015. The granular nature of exports is remarkable even when we zoom in to analyze export operations at the firm, destination and product level. The granularity of exports shapes the specialization of Spanish exports. If the top 10 firms were removed, Spain would lose revealed comparative advantage in 56% of 2-digit HS chapters, which account for 26% of all Spanish exports. This points out that superstars, along with country-level fundamentals, play a very important role in determining countries' trade patterns. Finally, we show that superstars can explain around one-third of the fluctuation of Spanish exports.

These results suggest that large differences across exporters should be a key ingredient of international trade models. The dominance of few firms might also qualify previous predictions on the effects of trade liberalization on firms entry and exit, aggregate productivity and welfare. In particular, as pointed out by Freund and Pierola (2015), small policy changes might have large effects if they alter the behavior of superstar firms.

References

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2016). Global firms. *NBER Working Paper No. 22727*, National Bureau of Economic Research.

Bernard, A. B., Jensen, J. B., and Schott, P. K. (2009). Importers, exporters and multinationals: A portrait of firms in the US that trade goods. In Dunne, T., Jense, J., and Roberts, M., editors, *Producer dynamics: New evidence from micro data*, pages 513–552. University of Chicago Press, Chicago.

Canals, C., Gabaix, X., Vilarrubia, J. M., and Weinstein, D. E. (2007). Trade patterns, trade balances and idiosyncratic shocks. *Banco de España Research Paper No. WP-0721*.

Costinot, A., Donaldson, D., and Komunjer, I. (2012). What goods do countries trade?

[20] Di Giovanni et al. (2014) analyze the contribution of the firm-specific component to the volatility of exports for all firms, not only superstars. They find that the contribution of this component is much larger than the contribution of the sector and destination component. In contrast, Canals et al. (2007) find that idiosyncratic shocks only explain 15% of total variation in aggregate exports in Japan.

A quantitative exploration of Ricardo's ideas. *The Review of Economic Studies*, 79(2):581–608.

de Lucio, J., Mínguez, R., Minondo, A., and Requena, F. (2011). The extensive and intensive margins of Spanish trade. *International Review of Applied Economics*, 25(5):615–631.

Di Giovanni, J. and Levchenko, A. A. (2012). Country size, international trade, and aggregate fluctuations in granular economies. *Journal of Political Economy*, 120(6):1083–1132.

Di Giovanni, J. and Levchenko, A. A. (2013). Firm entry, trade, and welfare in zipf's world. *Journal of International Economics*, 89(2):283 – 296.

Di Giovanni, J., Levchenko, A. A., and Mejean, I. (2014). Firms, destinations, and aggregate fluctuations. *Econometrica*, 82(4):1303–1340.

Eaton, J., Kortum, S. S., and Sotelo, S. (2012). International trade: Linking micro and macro. *NBER Working Paper No. 17864*, National Bureau of Economic Research.

Fernandes, A. M., Freund, C., Pierola, D., and Cebeci, T. (2015). Exporter behavior, country size and stage of development. *Policy Research Working Paper WPS7452*. The World Bank.

Freund, C. and Pierola, M. D. (2015). Export superstars. *Review of Economics and Statistics*, 97(5):1023–1032.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica*, 79(3):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.

Gaubert, C. and Itskhoki, O. (2016). Granular comparative advantage. *Available at http://www.princeton.edu/ itskhoki/papers/GranularCA.pdf*.

Gaulier, G. and Zignago, S. (2010). Baci: International trade database at the product-level. The 1994-2007 version. Working Papers 2010-23, CEPII.

Laursen, K. (2015). Revealed comparative advantage and the alternatives as measures of international specialization. *Eurasian Business Review*, 5(1):99–115.

Marin, D., Schymic, J., and Tscheke, J. (2015). Europe's export superstars - it's the organisation. *Bruegel Working Paper 2015/05*.

Mayer, T. and Ottaviano, G. I. (2008). The happy few: the internationalisation of European firms. *Intereconomics*, 43(3):135–148.

Minondo, A. (2013). Trading firms in the spanish service sector. *Revista de Economía Aplicada*, 21(63):5–28.

Van Beveren, I., Bernard, A. B., and Vandenbussche, H. (2012). Concording eu trade and production data over time. *NBER Working Paper No. 18604*, National Bureau of Economic Research.

Wagner, J. (2012). The German manufacturing sector is a granular economy. *Applied Economics Letters*, 19(17):1663–1665.