The granular and fundamental components of export specialization*

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

May 16, 2017

Abstract

Countries' export specialization patterns are often caused by the behavior of very few firms. We propose an easy-to-implement methodology to decompose export specialization into fundamental comparative advantage (a country-specific component) and granular comparative advantage (a firm-specific component). We apply this methodology to analyze export specialization across countries and across regions within a country. In the country-level analysis, we find that, on average, granular comparative advantage leads to export specialization in 29% of industries, which account for 47% of total exports. We also show that 60% of the variation in export specialization across countries is explained by granular comparative advantage. The contribution of firms to export specialization is more important across regions within a country than across countries.

JEL: F12, F14

Keywords: exports, fundamental comparative advantage, granular comparative advantage, European Union, Spain, regions, export superstars, firm-level data.

*We thank the Customs and Excise Department of the Spanish Revenue Agency (AEAT) for providing essential information for this paper. We also thank Carlos Llano, Antonio Navas, Jose M. Pavia and participants at the Universidad Autónoma de Madrid and the Universidad de Murcia's research seminars for helpful comments and suggestions. 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

The question of why countries export some goods and import others is central to positive trade theory (Jones and Neary, 1984). The Ricardian and Heckscher-Ohlin models of trade contend that countries export the goods in which they have a comparative advantage and import the goods in which they have a comparative disadvantage. In the Ricardian model comparative advantage emerges from differences in technology and in the Heckscher-Ohlin model from differences in factor endowments. Since these variables are determined at the country level, firms do not play any role in shaping export specialization.

However, there are well-known examples, such as Nokia in Finland, Samsung in Korea and Intel in Costa Rica, where a single firm has a large influence on a country's export specialization (Freund and Pierola, 2015). Another example is the case of Spain. This country had a strong export specialization in apparel in 2014. The share of this industry in Spanish exports in 2014 was 60% higher than in world exports.[1] However, only three firms accounted for 52% of all Spanish exports in this industry. Was Spanish export specialization in apparel explained by Ricardian and Heckscher-Ohlin comparative advantage, or by the performance of three outstanding firms?

In this paper, we propose an easy-to-implement methodology to decompose export specialization into fundamental comparative advantage (a country-specific component) and granular comparative advantage (a firm-specific component). We define export specialization as the value obtained from the following three steps. (i) For an origin-destination pair, we select two industries, and calculate the ratio of export values; (ii) for the same destination and industries, we select another country of origin and calculate the ratio of export values; (iii) we compute the ratio of relative export values. This calculation has a ratio of ratios structure. We show that in a Melitz-type trade model with a continuum of firms, whose productivity is Pareto distributed, export specialization is the product of Ricardian comparative advantage, denoted as fundamental comparative advantage, and the ratio of export cost ratios. If we select a country and a reference country with the same export costs ratio, the ratio of export cost ratios will cancel out, and export specialization will be determined by fundamental comparative advantage only. Moreover, export specialization will be equal to the ratio of exporter ratios.

In reality, countries do not have a continuum, but a finite number of firms. In this environment, a country's export specialization will be determined by fundamental comparative advantage and the luck that firms have in the productivity draw. For example, if a firm draws an outstanding productivity, it will dominate exports and define a country's

[1]Authors' calculations using the Comtrade database (available at http://comtrade.un.org). To measure apparel exports we add up the HS chapters 61 and 62.

export specialization. Granular comparative advantage captures the effect that drawing an outstanding productivity has on a country's export specialization. In a setting where the ratio of export costs are equal across countries, granular comparative advantage can be estimated as the difference between actual export specialization and fundamental comparative advantage.

To cancel out the ratio of export cost ratios at the country-level, we use intra-European Union (EU) export data, and investigate EU countries' export specialization relative to a reference EU country. Since there is mutual recognition in technical and safety standards across EU countries, there are no tariffs and quotas in intra-EU trade, and most EU countries use roads to transport their goods to other EU countries, it is reasonable to assume similar relative export costs across EU countries in their trade with other EU countries. We provide evidence to validate this assumption. To apply our simplification in a more stringent setting, we also investigate Spanish regions' export specialization relative to a reference Spanish region. We expect the similarities in the ratio of export costs to be even larger across regions, within a country, than across countries belonging to the same economic union.

Next, we show that the ratio of exporter ratios observed when the number of firms is finite provides an accurate estimation of the ratio of exporter ratios that we would observe if there was a continuum of firms. This allow us to estimate fundamental comparative advantage. Then, as explained before, we calculate granular comparative advantage as the difference between actual export specialization and fundamental comparative advantage.

The empirical analyses using country-level data show that, on average, granular comparative advantage leads to export specialization in 29% of the industries analyzed. These granular industries account for 47% of EU countries' exports to other EU countries. We also find that granular comparative advantage explains 60% of the variation in export specialization across EU countries, while fundamental comparative advantage explains the remaining 40% of the variation. The empirical analyses with regional data show that granularity plays a larger role in determining export specialization at the regional than at the country level. This result is in line with the fact that differences in the sources of fundamental comparative advantage are smaller across regions within a country than across countries. The information covered in the regional database allows us to explore additional questions. First, there is persistence in the industries that are classified as granular, and in the contribution of fundamental and granular comparative advantage to variations in export specialization over time. Second, industries where the top exporter is large are more likely to command a high granular comparative advantage; and, third, granularity is not more likely to emerge in industries with a large fundamental comparative advantage.

Our paper is related to the literature that analyzes the contribution of granular and fundamental comparative advantage to export specialization. A first attempt to estimate the contribution of these components is provided by Freund and Pierola (2015), who analyze whether countries' revealed comparative advantage in an industry would alter if the top exporters disappeared.[2] Granular comparative advantage is dominant if revealed comparative advantage disappears once the top exporters are removed; in contrast, fundamental comparative advantage is dominant if revealed comparative advantage remains even when the top firms are removed. The limitation of this methodology is that the behavior of the remaining exporters, once the top firms have disappeared, is not known. Besides, it only identifies whether an industry is granular or fundamental. Since they do not have a fundamental and granular comparative advantage measure, they cannot estimate the contribution of these components to the variation in export specialization across countries. To overcome these limitations, Gaubert and Itskhoki (2016) develop a general equilibrium model with a finite number of firms. They apply a simulated method of moments to obtain the values of the parameters that minimize the differences between the moments generated by the model and the actual moments. These values are then used to calculate fundamental and granular comparative advantage in French industries. Our paper, while maintaining the features of a general equilibrium model, contributes to this literature by offering an alternative, easy-to-implement methodology that enables identification of granular industries and measurement of the contribution of granular and fundamental comparative advantage to the variation in export specialization across countries and across regions within a country.[3] By applying our methodology, to the best of our knowledge, we are the first authors to compare the contribution of granular and comparative advantage to export specialization across countries.

Our paper is also related with the literature that has analyzed export specialization at the regional level (Courant and Deardorff, 1992; Coşar and Fajgelbaum, 2016). As far as we know, for the first time in the literature, we present data on the concentration of exports by firm at the regional level. We show that in some regions the top exporter might account for almost 35% of exports. Another novelty is that our approach allows estimation of the contribution of granular and fundamental comparative advantage to differences in export specialization across regions. We find that most of the differences in export specialization are explained by granular comparative advantage.

The rest of the paper is organized as follows. Section 2 analyzes the variables that determine export specialization in a Melitz-type trade model, where productivity is Pareto distributed. Section 3 explains our methodology to estimate fundamental and granular comparative advantage, and applies this to data from countries and regions within a country. Section 4 concludes.

[2] De Lucio et al. (2017) also apply this methodology for Spanish exports.

[3] We could not apply the methodology proposed by Gaubert and Itskhoki (2016) because they use moments that demand information on domestic sales that we do not have.

2 Export specialization in a Melitz-type model where firms' productivity is Pareto distributed

We begin by decomposing exports into the number of exporters, the extensive margin, and the average exports per firm, the intensive margin:

$$X_{ijk} = N_{ijk}\overline{x}_{ijk} \tag{1}$$

where N_{ijk} is the number of firms located in country i that export industry k varieties to country j, and $\overline{x}_{ijk}$ is the average exports per firm.

To investigate the determinants of the extensive and intensive margins of exports, we frame our analysis in a Melitz-type heterogeneous firms trade model (Melitz, 2003). Firms produce horizontally-differentiated varieties within an industry with monopolistic competition, labor is the only production factor, and preferences of a representative consumer are given by a constant elasticity of substitution (CES) utility function. Firms are heterogeneous in productivity and face fixed and variable export costs. Following Chaney (2008), the potential number of entrants is fixed, but large enough for there to be a continuum of firms.

As is standard in the literature, we assume that firms' productivity is Pareto distributed. Arkolakis et al. (2012) provide three reasons to explain the popularity of this distribution function: (i) it is easy to treat analytically; (ii) it provides a reasonable approximation of the distribution of firm sales and exports; (iii) from a theoretical point of view, the Pareto distribution can be the outcome of simple stochastic processes of firm-level growth, entry and exit.

As shown in Appendix A, if productivity is Pareto distributed, the intensive margin of exports is determined by

$$\overline{x}_{ijk} = \Big(\frac{\theta\sigma}{\theta - \sigma + 1)}\Big)F_{ijk} \tag{2}$$

where σ is the elasticity of substitution and F_{ijk} is the fixed cost of exporting k industry varieties from country i to country j. The shape parameter θ measures the heterogeneity of the productivity distribution, with higher values meaning less heterogeneity.

As explained in Appendix A, the extensive margin of exports is determined by the following expression:

$$N_{ijk} = M_i\Big(\frac{z_{ijk}}{\varphi_{ik}}\Big)^{-\theta} \tag{3}$$

where M_i is the exogenous mass of firms that can potentially enter any industry in country i; $\underline{z}_{ijk}$ is the threshold productivity that firms in country i should reach in order to obtain profits from exporting k industry varieties to country j; and, φ_{ik} is the minimum productivity that firms can draw in country i and industry k. Following Costinot et al. (2012), we denote this last parameter as the fundamental productivity of country i in industry k. According to (3), the number of exporters will be larger the lower the threshold productivity to export, and the larger the fundamental productivity and the heterogeneity of the distribution of productivity.

As shown in Appendix A, the export-threshold is determined by the following expression:

$$\underline{z}_{ijk} = \left(\frac{F_{ijk}}{\mu\beta_{jk}Y_j}\right)^{(1/\sigma-1)}\left(\frac{w_i\tau_{ijk}}{P_{jk}}\right) \tag{4}$$

where $\mu = (\sigma-1)^{\sigma-1}\sigma^{-\sigma}$; β_{jk} is the share of income that country j devotes to consuming industry k varieties; Y_j is the income of country j; w_i is the wage in country i; τ_{ijk} is an iceberg-type trade cost, denoting the units of an industry k variety that should be sent from country i to ensure that one unit arrives in country j; finally, P_{jk} is the price index of industry k varieties in country j.

Substituting (2), (3) and (4) in (1) we can express the total value of exports as

$$X_{ijk} = \varphi_{ik}^{\theta}F_{ijk}^{\frac{\sigma-\theta-1}{\sigma-1}}\tau_{ijk}^{-\theta}M_i(\mu\beta_{jk}Y_j)^{\frac{\theta}{\sigma-1}}\left(\frac{w_i}{P_{jk}}\right)^{-\theta} \tag{5}$$

We define country i export specialization in industry k and destination j as

$$XS_{ijk} = \frac{X_{ijk}/X_{ijk'}}{X_{i'jk}/X_{i'jk'}} \tag{6}$$

where k' is the reference industry and i' is the reference country. Country i export specialization in industry k and destination j is measured as exports of industry k, relative to exports in a reference industry k', divided by the same ratio in a reference country i'.

If we substitute (5) in (6), the variables M_i, μ, β_{jk}, Y_j, w_i, and P_{jk} cancel out, leaving the expression

$$XS_{ijk} = \left(\frac{\varphi_{ik}/\varphi_{ik'}}{\varphi_{i'k}/\varphi_{i'k'}}\right)^{\theta}\left(\frac{F_{ijk}/F_{ijk'}}{F_{i'jk}/F_{i'jk'}}\right)^{\frac{\sigma-\theta-1}{\sigma-1}}\left(\frac{\tau_{ijk}/\tau_{ijk'}}{\tau_{i'jk}/\tau_{i'jk'}}\right)^{-\theta} \tag{7}$$

Note that the first ratio of ratios on the right-hand side of (7) is the Ricardian comparative advantage of country i in industry k. We denote the Ricardian comparative advantage as fundamental comparative advantage. According to (7), when there is a continuum of firms, export specialization is the product of fundamental comparative advantage, the ratio of fixed export costs ratios, and the ratio of variable export costs ratios.

If we substitute (3) in (7) we can also express export specialization as

$$XS_{ijk} = \Big(\frac{N_{ijk}/N_{ijk'}}{N_{i'k}/N_{i'k'}}\Big)\Big(\frac{F_{ijk}/F_{ijk'}}{F_{i'jk}/F_{i'jk'}}\Big)\Big(\frac{\tau_{ijk}/\tau_{ijk'}}{\tau_{i'jk}/\tau_{i'jk'}}\Big) \tag{8}$$

According to this latter equation, export specialization is the product of the ratio of exporter ratios, the ratio of fixed export costs ratios and the ratio of variable export costs ratios. In the next section, we will use (7) and (8) to estimate the fundamental comparative advantage component of export specialization.

3 Empirical analyses

We divide this section into three parts. First, we explain the methodology to estimate fundamental and granular comparative advantage. Second, we apply this methodology to analyze the contribution of fundamental and granular comparative advantage to the variation in export specialization across countries. Finally, we apply the methodology to investigate the contribution of fundamental and granular comparative advantage to the variation in export specialization across regions within a country.

3.1 A methodology to estimate the fundamental and granular components of export specialization

If we identified a country i and a country i' where $F_{ijk}/F_{ijk'} = F_{i'jk}/F_{i'jk'}$ and $\tau_{ijk}/\tau_{ijk'} = \tau_{i'jk}/\tau_{i'jk'}$, using (7) export specialization would be determined by fundamental comparative advantage, raised to a distribution parameter, only

$$XS_{ijk} = \Big(\frac{\varphi_{ik}/\varphi_{ik'}}{\varphi_{i'k}/\varphi_{i'k'}}\Big)^{\theta} \tag{9}$$

Using (8), export specialization could also be expressed as

$$XS_{ijk} = \Big(\frac{N_{ijk}/N_{ijk'}}{N_{i'k}/N_{i'k'}}\Big) \tag{10}$$

To apply this simplification, for the country-level analysis we use intra-EU exports, and investigate EU countries' export specialization relative to a reference EU country. As argued by Helpman et al. (2008), fixed export costs combine the costs that exporters face in their country (e.g. the costs of drafting a contract for a foreign delegate) and in the destination country (e.g. the legal costs of opening a delegation). There are differences in regulatory and legal costs across EU countries.[4] However, since we measure fixed export in an industry, relative to another industry, these country-level differences should cancel out. Since there is mutual recognition among EU countries, there are no fixed costs of adapting products to meet the technical and safety standards of destination countries. For some bilateral flows, fixed (and variable) export costs are lower because partners speak the same language. However, we expect the cost-reduction effect of a shared language to be similar across sectors, and to cancel out when measuring relative export costs. Variable export costs combine transport and other trade barriers, such as communication costs and tariffs. Transport costs, as explained by Combes and Lafourcade (2005), depend on itinerary, transport mode and commodity. Most trade among EU countries uses roads as the mode of transport, so it is reasonable to assume that commodities will follow the same itinerary from the EU country of origin to the EU country of destination. In addition, since there are no tariffs or quotas in intra-EU trade, we argue that it is also reasonable to assume similar ratios of variable export costs across EU countries. To validate our assumption, we proxy export costs with the Cost Insurance Freight/Free on Board (CIF/FOB) value ratio, and calculate the ratio of export cost ratios for all intra-EU export flows. As expected, we find that the median ratio of trade cost ratios has the value of one. To abide with the assumption of equal relative trade cost, we remove from the sample the EU countries that lead to a ratio of trade cost ratios that deviates significantly from 1.

To abide more tightly by the simplification of equal export costs ratios, we also investigate the export specialization of Spanish peninsular regions relative to a reference Spanish peninsular region in their trade with France. Regulatory and legal costs are very similar across Spanish regions, and all face the same relative variable export costs when exporting to France. Hence, we expect the similarities in relative fixed and variable export costs to be more pronounced across Spanish regions than across EU countries.

According to (9) and (10), if relative fixed and variable export costs are the same, export specialization is determined by fundamental comparative advantage only, and is equal to the ratio of exporter ratios. A key contribution of our paper is to show that even when the number of firms is finite, the ratio of exporter ratios still provides an accurate estimate of fundamental comparative advantage. When the number of draws is finite, the distribution of productivities might differ from the one predicted by the density function with a continuum of firms. However, in this methodology, we only need to know the number of exporters. This is equivalent to reducing the productivity levels that firms might draw to two: a productivity below the export-threshold, and

[4]For example, according to the World Bank's Doing Business 2017 report, the ease of doing business in Denmark, the United Kingdom or Sweden was much higher than in Greece or Italy. Available at: http://www.doingbusiness.org/~/media/WBG/DoingBusiness/Documents/Annual-Reports/English/DB17-Report.pdf

a productivity above the export-threshold. In this case, even with a finite, and small, number of draws, the distribution of firms between exporters and non-exporters would be very similar to the distribution that we would observe if the number of draws was infinite. Hence, the observed ratio of exporter ratios provides an accurate approximation of the ratio of exporter ratios that we would have observed in the continuum case.

To support our argument, we draw on Eaton et al. (2012), who show that if the number of draws is finite, the number of firms in country i that export k industry varieties to country j is the realization of a random variable that follows a Poisson distribution, with parameter $\lambda = M_i(\underline{z}_{ijk}/\varphi_{ik})^{-\theta}$. Note that in a Poisson distribution the expected value of the random variable is λ. Hence, the expected exporters in a finite number of draws scenario is the same as the exporters in a scenario with an infinite number of draws. In a Poisson distribution, the standard deviation of the random variable is $\sqrt{\lambda}$. To measure the extent to which a realization might differ from the expected value in each of the four elements that compose the ratio of exporter ratios, we calculate the number of exporters' coefficient of variation:

$$cvN_{ijk} = \frac{(\underline{z}_{ijk}/\varphi_{ik})^{\theta/2}}{\sqrt{M_i}} \tag{11}$$

We give values to the variables in (11) to measure the coefficient of variation. The ratio in the numerator measures the minimum productivity that firms in country i need to reach in order to export industry k varieties to country j, relative to the fundamental productivity of firms in country i and industry k. We can approximate this ratio with the exporters' labor productivity premium estimated by the empirical literature. For example, Bernard et al. (2007a) report that value-added per worker is 11 percent larger in exporters than non-exporters in the US, once industry effects are controlled for. Following Eaton et al. (2012), we take $\theta = 5$. Even for a very small number of draws[5], $M_i = 100$, the coefficient of variation is very low, $cvN_{ijk} = 0.13$.[6] This conclusion is in line with Minondo (2017), who compares the share of expert chess players across countries predicted by a model with a continuum of players and a model with a finite number of players. Using a simulated method of moments, he shows that, for moderate levels of expertise equivalent to a low $\underline{z}_{ijk}/\varphi_{ik}$ ratio, the continuum and discrete models predict very similar percentages.

As shown in Appendix B, the coefficient of variation of the ratio of exporter ratios is determined by a more complex expression than (11). We use random numbers generated by a Poisson distribution with different λ parameters to measure the coefficient of variation in alternate scenarios. As shown in Table A1, even in very stringent scenarios, the

[5]For example, Gaubert and Itskhoki (2016) use 8,000 draws for small French sectors.

[6]A distribution with a coefficient of variation lower than 1 is considered a low-variance distribution.

coefficient of variation remains low.

According to (9) and (10), there is a one-to-one relationship between export specialization and fundamental comparative advantage, and a one-to-one relationship between this latter variable and the ratio of exporter ratios. We define granular comparative advantage as the difference between actual export specialization and the export specialization that we would find if there was a continuum of firms, which, in turn, would be equal to the ratio of exporter ratios. It is important to stress that the difference between actual export specialization and the ratio of exporter ratios can be positive or negative. A positive difference happens when a firm draws an outstanding productivity, driving the industry's intensive margin above the average. As highlighted in the introduction, the standard concept of granularity is related with this positive difference. However, granularity can be negative too. This case, which is less associated with granularity, captures situations where the productivity drawn by firms is lower than expected, driving the intensive margin below the average.

We should stress that, empirically, granular comparative advantage is calculated as a residual of the difference between actual export specialization and the ratio of exporter ratios. Hence, our measure of granular comparative advantage might also capture differences that are not purely granular, such as the error term between the true ratio of exporter ratios and the one we observe, small differences in relative variable and fixed export costs that remain in our samples, or a distribution of productivity across firms that might depart from Pareto.[7] Since fundamental comparative advantage captures Ricardian comparative advantage only, granular comparative advantage might also encompass other sources of comparative advantage, such as differences in factor endowments.[8] At the country level, we expect the effect of this alternative source of comparative advantage to be attenuated, since we select countries with similar relative factor endowments. We expect the effect to be very small when applying our methodology at the regional level. Hence, we should consider our methodology to be a fairly accurate approximation for identifying fundamental and granular comparative advantage.

[7] Some papers explore whether other probability functions, such as the log-normal, approximate the distribution of productivity across firms more accurately (Head et al., 2014; Fernandes et al., 2015b; Bas et al., 2016). These papers show that if productivity is distributed log-normal, there is a positive relationship between the extensive and the intensive margin of exports. Hence, in a log-normality scenario, the number of exporters in industry k relative to industry k' would underestimate the exports in industry k relative to industry k'. We expect the ratio-of-ratios structure of our indicators to attenuate this undervaluation. In any case, in order to account for log-normality, we should qualify the fundamental comparative estimates obtained with our methodology as lower bound estimates. Nigai (2017) shows that Pareto provides a good fit for the upper-right tail of the distribution, whereas the log-normal provides a good fit for the rest of the distribution.

[8] For example, Bernard et al. (2007b) embed a heterogeneous-firms model in a Heckscher-Ohlin framework.

3.2 Empirical analyses with country-level data

Our first empirical analysis is carried out using EU countries' exports to other EU countries. Data on the number of exporters and value of exports per industry is obtained from the OECD-Eurostat Trade by Enterprise Characteristics Database (Araújo and Gonnard, 2011; Eurostat, 2016).[9] For each EU country, the database provides information on the number of exporters and the value of exports to EU countries for 21 manufacturing industries over the 2008-2013 period.[10]

To evaluate the validity of the equal relative export costs assumption, we draw data from Eurostat's database on the value and quantity of exports and imports between each EU country for the 96 industries included in the HS 2-digit classification.[11] For each bilateral flow, we compare the value and quantity of exports reported by the EU country of origin[12], measured FOB, with the mirror value and quantity of imports reported by the EU country of destination, measured CIF. If all statistical agencies followed the same registration methods, and there were no errors recording trade flows, the quantity reported by the exporter should equal the mirror quantity reported by the importer; and the import value ought to be larger than the mirror export value. If these conditions were met, the CIF/FOB ratio would be a good proxy for export costs. However, for many flows these conditions are not met. The researcher might find that, for some transactions, the FOB value is larger than the mirror CIF value, and might observe substantial discrepancies between the export quantity and the mirror import quantity (Hummels and Lugovskyy, 2006). To calculate consistent CIF/FOB ratios, we follow the methodology proposed by Guillaume et al. (2008). First, we select the flows where the exported quantity/imported quantity ratio is in the [0.9-1.1] range. Only 15% of the flows in our dataset fall in that range. Second, we calculate a unit value-based CIF/FOB ratio:

$$CIFu/FOBu = \frac{p_{ij}^M q_{ij}^M / q_{ij}^M}{p_{ij}^X q_{ij}^X / q_{ij}^X} \tag{12}$$

where q_{ij}^M is the quantity that country j imports from country i, and q_{ij}^X is the quantity that country i exports to country j; p_{ij}^M and p_{ij}^X are the price of imports and exports, respectively. The numerator provides the unit value of imports, measured CIF, and the denominator measures the unit value of exports, measured FOB. We only select trade

[9] This database is available at http://stats.oecd.org

[10] Since data for tobacco and refined petroleum industries are confidential in many countries, we exclude them from the analysis. The industries included in the sample are: Food products, beverages, textiles, apparel, leather, wood, paper, printing, chemicals, pharmaceuticals, rubber, other non-metallic mineral products, basic metals, metal products, computers and electronics, electrical equipment, machinery, motor vehicles, other transport equipment, furniture and other manufacturing.

[11] Data are available at http://ec.europa.eu/eurostat. We use data for the year 2008.

[12] Due to their very small size, we exclude Cyprus, Luxembourg and Malta from the sample.

flows whose CIFu/FOBu ratio is in the [1-2] range. This additional condition reduces the sample to 8% of the original observations.

We calculate the ratio of trade cost ratios for all origin, destination and industry combinations. The median ratio of ratios is 0.992 and the mean is 1.011; 80% of the ratio of ratios are in the 0.76-1.29 range. These figures confirm that it is reasonable to assume similar trade cost ratios across EU countries.[13] To strengthen the validity of the assumption of equal relative export costs, we analyze whether some EU countries have a large number of ratios of trade cost ratios that significantly differ from the median value. In particular, for each EU country, we identify the ratios of trade cost ratios that are below 0.9 or above 1.1. For each country, we calculate the trade value captured in those outlier ratios as a share of the total trade value captured in the ratios in which the country participates. We find that there are 10 countries whose trade in outlier ratios represents more than 50% of their trade. We decide to remove those countries from the sample.[14]

Table A2 in the Appendix presents summarized statistics for the 14 EU countries included in the sample. Most countries provide data for the whole 2008-2013 period, although there are a few countries, such as Belgium, that provide data for three years only. For each country, we calculate the total number of exporters and the total value of exports to EU countries as the average over the period of analysis. Italy is the country with the highest number of exporters, followed by Germany and Poland. Germany is the largest exporter to other EU countries, followed by Italy, France and Belgium. There is a positive correlation between the number of exporters and GDP (0.75). This result is in line with previous studies that have analyzed the relationship between countries' economic size and the number of exporters (Fernandes et al., 2015a). We also find a positive correlation between GDP and exports (0.95), a result which is in line with numerous previous studies on the gravity of trade (Head and Mayer, 2014). Finally, the table shows that most manufacturing exports to other EU countries are transported by road.[15]

To calculate the ratio of exporter ratios, and the ratio of export value ratios, we need a reference country and a reference industry. To maximize observations, we should select as a reference the country that exports in most industries, and the industry with the highest number of exporting countries. However, there are several countries and industries that meet these criteria. Since some empirical calculations might be sensitive to a specific reference country and industry, we calculate the ratio of exporter ratios and the ratio of

[13] However, it is important to stress that the country-level database has an aggregate destination, EU countries, and an aggregate industry classification. These aggregations might weaken the validity of the assumption of similar relative export costs, due to heterogeneity in the destination portfolio and of the products that are exported within each industry across EU countries.

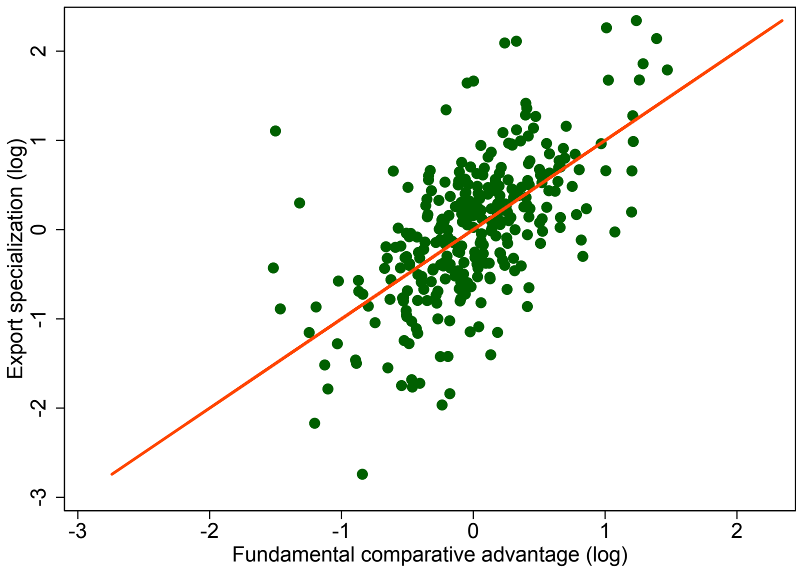
[14] The EU countries removed from the sample are Bulgaria, the Czech Republic, Denmark, Estonia, Finland, Greece, Ireland, Latvia, Sweden and the United Kingdom.

[15] We get these data from Eurostat's international trade database. Data correspond to the year 2013.

export value ratios for all possible reference country+industry combinations. Then, we calculate the average of each ratio for each country and industry. Hence, we measure EU countries' export specialization relative to an average EU country and industry.[16]

Since the ratio of export value ratios and the ratio of exporter ratios are not bounded from above, they might take outlier values. To attenuate the effect of outliers, we transform export specialization and the ratio of exporter ratios into log values. We calculate granular comparative advantage as the difference between the log of actual export specialization and the log of the ratio of exporter ratios.

Figure 1: (Log) Export specialization vs. (log) fundamental comparative advantage (Average 2008-2013)



Source: Authors' calculations using the OECD-Eurostat Trade by Enterprise Characteristics database.

Figure 1 presents a scatter diagram of (log) export specialization and (log) fundamental comparative advantage. There is a positive correlation between both variables: the larger the fundamental comparative advantage, the larger the export specialization. If export specialization was explained by the fundamental comparative advantage only, all dots would lie on the 45° line. However, we observe that dots scatter around the 45° line. Our methodology ensures that these deviations mostly capture granular effects. The dots above the 45° line are country+industry combinations where granular comparative advantage is positive, whereas the dots below the 45° line are country+industry combinations where granular comparative advantage is negative.

We first identify the industries in which EU countries reveal a comparative advantage

[16]This methodology also strengthens the assumption of equal relative export costs, since the median ratio of trade cost ratios is 1.

(export specialization >1). Among them, we define as granular the industries where granular comparative advantage is larger than fundamental comparative advantage. For each country, we calculate the percentage of granular industries, the share of granular industries among the industries with a revealed comparative advantage, and the share of granular industries in total exports. Table 1 presents these calculations for the 14 EU countries included in our sample. On average, 29% of industries are granular, they represent 56% of industries with a revealed comparative advantage and account for 47% of exports.

Table 1: Granular industries by country (Average 2008-2013)

Country	% of granular industries	Share of granular in XS >1 industries	% of granular exports
Average country	29	56	47
Austria	29	60	36
Belgium	24	56	44
France	14	30	30
Germany	14	30	38
Hungary	38	80	71
Italy	29	50	29
Lithuania	33	70	63
Netherlands	38	62	70
Poland	33	70	64
Portugal	19	40	30
Romania	33	70	57
Slovakia	43	69	66
Slovenia	29	46	27
Spain	29	55	47

Source: Authors' calculations using the OECD-Eurostat Trade by Enterprise Characteristics database. Note: XS=Export specialization. Among industries with export specialization >1, granular industries are defined as those where granular comparative advantage is larger than fundamental comparative advantage.

The highest percentage of granular industries, 43%, is found in Slovakia, and the lowest percentage, 14%, in France and Germany. In Hungary, granular industries represent 80% of the industries in which this country has a revealed comparative advantage. The percentage drops to 30% in France and Germany. Exports generated in granular industries represent 71% of Hungarian exports, but only 27% of Slovenian exports. There is a negative correlation between GDP and the percentage of exports in granular industries (-0.43). However, there are countries, such as Slovenia or Portugal, where the share of granular industries in total exports is lower than in larger countries, such as Germany or Spain. This suggests that, along with country size, what industries are granular also determines the weight of granular industries in total exports.

Next, we analyze the contribution of fundamental comparative advantage (FCA) and granular comparative advantage (GCA) to explain the differences in export specialization across countries. To perform this analysis, we use a regression-based decomposition. We regress each comparative advantage on export specialization and a constant. Specifically,

$$\begin{aligned} FCA_{ijk} &= \ln\left(\frac{N_{ijk}/N_{ijk'}}{N_{i'jk}/N_{i'jk'}}\right) = \alpha + \beta_1 \ln XS_{ijk} \\ GCA_{ijk} &= \ln XS_{ijk} - FCA_{ijk} = \alpha + \beta_2 \ln XS_{ijk} \end{aligned} \tag{13}$$

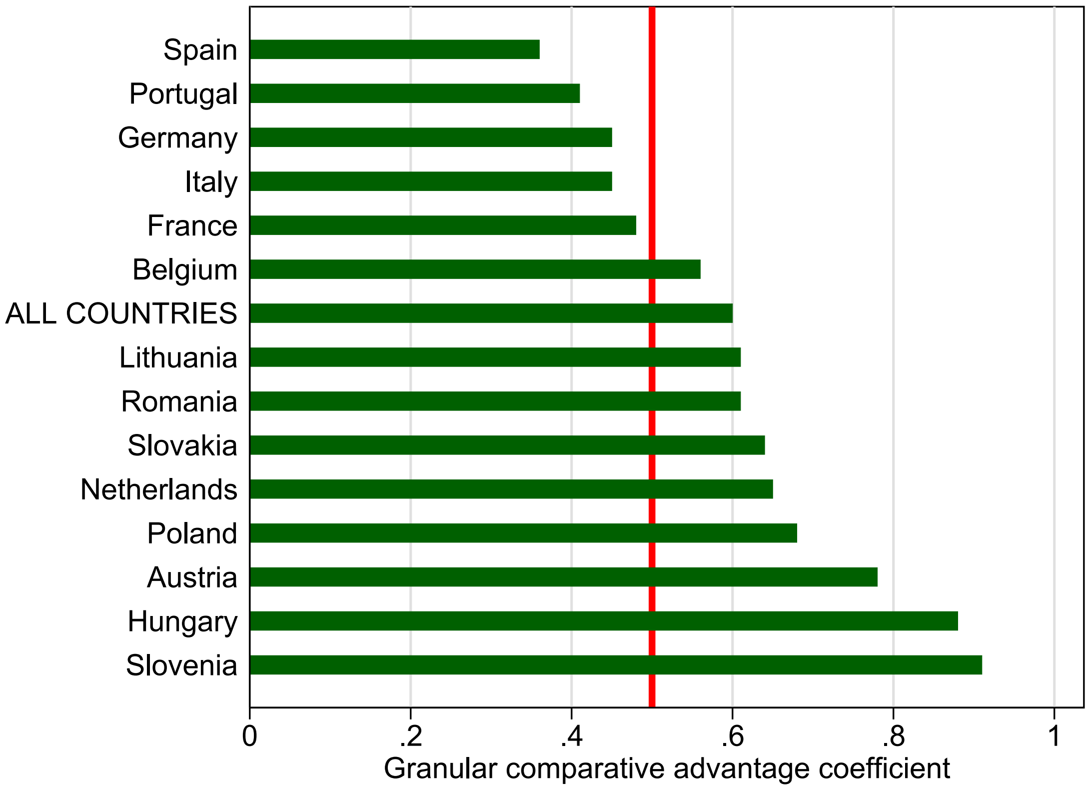
Figure 2 presents the results of the regression-based decomposition for granular comparative advantage. First, we perform the decomposition by pooling all observations; next, we carry out country-specific decompositions. When we pool all observations, 60% of variations in export specialization across countries are explained by granular comparative advantage and 40% by fundamental comparative advantage. These results show that, on average, granular comparative advantage plays a larger role than fundamental comparative advantage in explaining the differences in export specialization across countries.[17]

Next, we estimate the contribution of fundamental and granular comparative advantage to the variation in export specialization within each country. To do so, we run a separate regression for each country. In 9 out of the 14 countries included in the sample, the contribution of granular comparative advantage is larger than the contribution of fundamental comparative advantage, while the opposite is the case in the other five countries: France, Germany, Italy, Portugal and Spain. The highest contribution of granular comparative advantage is in Slovenia (91%), followed by Hungary (88%) and Austria (78%). The countries with the highest contribution of fundamental comparative advantage are Spain (64%), Portugal (59%) and Germany (55%). We find a positive correlation between GDP and the contribution of fundamental comparative advantage (0.58), which might be explained by the relative lower influence of large exporters in high GDP countries.

Gaubert and Itskhoki (2016) estimate the contribution of granular and comparative advantage to differences in export specialization across French industries. They find that 70% of the variation in export specialization is due to fundamental comparative advantage and 30% is due to granular comparative advantage. Their estimate of the fundamental comparative advantage contribution, 70%, is larger than our estimate, 52%. Gaubert and Itskhoki (2016) analyze France's export specialization vis a vis the rest of the world, while we analyze it vis a vis an average EU country. Since differences in the sources of comparative advantage are larger across the world than across the EU, it is reasonable to expect a higher contribution for fundamental comparative advantage when export specialization is measured relative to the former than to the latter.

[17] When we estimate (13) with industry fixed effects or country fixed effects, the results do not change.

Figure 2: Contribution of granular comparative advantage to variation in export specialization. Regression-based decomposition (Average 2008-2013)



Source: Authors' estimations using the OECD-Eurostat Trade by Enterprise Characteristics database. Note: To calculate the contribution of granular comparative advantage we regress granular comparative advantage on export specialization. ALL COUNTRIES' regression pools the observations from all countries and industries.

We use the ratio of exporter ratios to estimate fundamental comparative advantage. The empirical literature shows that many firms export one year and cease to export the following year, which might introduce noise to our estimate.[18] This literature also suggests that firms with more employees are more likely to be regular exporters. We test whether our main results are altered if we select exporters with 10 or more employees only. Due to the absence of data, the sample is reduced to 9 countries. Table A3 presents information on granular industries. The percentage of granular industries, their share in industries where export specialization > 1, and their share in total exports are very similar to those found in the baseline analysis. Country-level results are also very similar. Figure A1 presents the regression decomposition results. When all countries and industries are pooled, the contribution of fundamental comparative advantage rises to 43%, and the contribution of granular comparative advantage declines to 57%. In any

[18]See, among others, Görg et al. (2012), Cadot et al. (2013), Esteve-Pérez et al. (2013) and Albornoz et al. (2016).

case, these percentages are similar to those obtained in the baseline analysis.

To sum up, our analyses show that granular comparative advantage plays a very important role in shaping a EU country's export specialization relative to an average EU country. On average, granularity defines export specialization in 29% of industries and granular industries account for 47% of total manufacturing exports. Moreover, granular comparative advantage explains 60% of the differences in export specialization across countries.

3.3 Empirical analyses with regional data

We use Spanish regional data to perform the second set of empirical analyses. The use of regional data is interesting for several reasons. First, as demanded by our methodology, regional data enables us to calculate fundamental and granular comparative advantage using exporter and export value data for a single destination. Second, regional data allow us to abide more tightly by the assumption of similar relative export costs. Choosing a single destination reduces the heterogeneity in the destination-portfolio that could be present in the country-level database. Moreover, the regional database has a more disaggregated industry classification, which reduces the heterogeneity in the product-portfolio that might be present in the country-level database. Third, the regional database allows us to reduce the noise in the number of exporters, a key variable in our calculations. Fourth, our regional database provides information over a long period, enabling dynamic analyses. Fifth, differences in relative factor endowments are smaller across Spanish regions than across countries. Hence, granular comparative advantage is less likely to capture Heckscher-Ohlin-based comparative advantage.[19]

Our data are obtained from the Customs Database, collected by the Customs and Excise Department of the Spanish Tax Agency, which covers all export transactions in Spain. For each transaction, we know a firm's identification code given by Customs, the product at the 8-digit Combined Nomenclature (CN) classification, the destination of the export transaction, the free-on-board (FOB) value in euros of the transaction, and its regional origin. Our database covers the 1997-2015 period. For the baseline analyses, we use 2014 data. To analyze persistence, we compare the baseline results with the results obtained using data from 1998. Since a firm might have plants in different regions, we identify exporters as the combination of the firm identification code and the region in which the export operation originates.

Spain is divided into 17 regions (Eurostat's NUTS II classification), which are shown

[19] As shown by Requena et al. (2008), a strict Heckscker-Ohlin-Vanek (HOV) empirical model, assuming identical technologies and preferences, does not fit well with Spanish regions' factor content of trade. In a follow-up paper, Artal-Tur et al. (2011) show that an HOV model fits better when technological (Ricardian) differences across regions are allowed.

in Map A1 in the Appendix. Due to their special geographic features, for the empirical analyses, we do not include the two regions located in Africa (Ceuta and Melilla), and we remove from the sample the two island regions (Balearic Islands and Canary Islands). To perform the empirical calculations we collapse exports at the HS 2-digit level, which distinguishes 96 different products, denoted as chapters.[20] A more disaggregated classification would allow a more detailed export specialization range, but at the cost of finding no exporters in a larger number of regions, which precludes the estimation of the ratio of exporter ratios. We achieve the balance between the level of detail of export specialization and a positive number of exporters per region at the HS 2-digit disaggregation level.

Columns 1 and 2 in Table 2 provide information on Spanish regions' share of total exports and number of exporters.[21] Almost half of Spanish exporters are located in Catalonia and Madrid. The next regions in the ranking are Valencia, 13%, and Andalusia, 11%. The region with the highest export value is Catalonia, which accounts for 25% of total exports, followed by Madrid, 12%, Andalusia, 11% and Valencia, 11%. It is interesting to observe that Madrid's share of total exports is much lower than that for total exporters. Many multi-plant exporters have a plant in Madrid, the capital of Spain. However, many export transactions are shipped from other Spanish regions.

Following Gaubert and Itskhoki (2016), as a first proxy of granular comparative advantage, Table 2 also presents the share of the top exporter and the share of the top 5 exporters in regional exports. There is a major variation in the share of the top exporter across regions. For example, in Castile and León the top exporter represents 35% of total regional exports, whereas in Madrid, the top exporter only represents 5% of all exports. The average is 15%, with a standard deviation of 10%. At the bottom of the table, we also present the share of the top exporter for Spain: 2%.[22] At the regional level, the correlation between exports and the share of the top exporter is -0.22. The differences across regions are still sizable for the share of the top 5 exporters. The average is 33%, with a standard deviation of 12%. The regions with the highest shares are Castile and León, 54%, Asturias, 47% and Navarre, 47%. The lowest percentages are found in Catalonia, 16%, Castile-La Mancha, 21% and Madrid, 21%. The share of the top 5 exporters in Spanish exports is 10%. These figures suggest that in some regions, a small number of firms dominate exports.

As in country-level analyses, for each region, we calculate the ratio of exporter ratios and the ratio of export value ratios for all reference regions and reference chapters; then, we take the average of those ratios. Hence, we measure the export specialization of a

[20] There is no chapter 77, and there are no data for chapters 98 and 99.

[21] The shares are calculated over the total number of exporters and the value of exports of the regions included in the sample.

[22] The Spanish figure is calculated with data from all regions.

Table 2: Distribution of exporters, export values and share of top exporters in Spanish regions, 2014

Region	% of total exporters	% of total exports	Top firm's share of regional exports	Top 5 firms' share of regional exports
Andalusia	11	11	14	31
Aragon	2	4	28	46
Asturias	1	2	18	47
Basque Country	8	10	9	24
Cantabria	1	1	10	34
Castile and León	3	5	35	54
Castile-La Mancha	3	2	9	21
Catalonia	27	25	9	16
Extremadura	1	1	6	23
Galicia	3	8	16	45
Madrid	21	12	5	21
Murcia	3	4	21	42
Navarre	1	3	32	47
Rioja	1	1	7	23
Valencia	13	11	8	24
Spain (total)			2	10

Source: Authors' calculations using the Customs database. Note: The regional number of exporters and exports shares are calculated over the total number of exporters and the value of exports of the regions included in the sample. The top 1 and top 5 figures for Spain are calculated using data from all regions.

Spanish region relative to an average Spanish region and industry. To reduce noise in the number of exporters, we remove small and occasional exporters. First, we exclude exporters whose total annual export operations in an HS 2-digit chapter are below 6,000 euros.[23] Next, we select firms that export a chapter for three consecutive years. Since our reference year is 2014, for each chapter, we select firms that export in 2013, 2014 and 2015. Finally, since our methodology demands export specialization to be measured at a given destination, we select France, the most important destination for Spanish exports.

We identify Spanish regions' granular industries, their share in the industries with a revealed comparative advantage, and their share in regional exports. First, we perform the analysis for the year 2014. To analyze persistence, we also carry out the analysis for the year 1998. Table 3 presents the results of these analyses. In 2014, on average, 37% of chapters of a Spanish region were granular, they represented 70% of the chapters with a revealed comparative advantage, and accounted for 62% of total regional exports. The figures for the country-level analysis were 29%, 56% and 47%, respectively. These figures

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

suggest that the weight of granular industries is larger across regions within a country than across countries. This result is in line with our expectations, since differences in the sources of comparative advantage are smaller across regions within a country than across countries. We observe some differences across regions, especially regarding the share of granular exports. There are some regions, such as Asturias, Castile and León and Galicia, where more than 90% of exports are accounted for by granular industries. In contrast, granular industries only represent 25% of exports in Valencia.

The results for 1998 are very similar to those for 2014. This suggests that there is persistence in the contribution of granular industries to regional exports. However, amid persistence, there are some regions, such as Aragon, Asturias and Murcia, that experience large swings in the share of granular industries in total exports.[24] To analyze persistence more accurately, we calculate the industries that were granular in 1998 and 2014, as a share of granular industries in 1998. We find that 58% of industries were granular in both years. If we restrict the analysis to industries that had a revealed comparative advantage in 1998 and 2014, persistence rises to 84%. These percentages confirm that there is persistence in the industries that are granular over time.

In our second empirical analysis, we investigate the variables that are positively correlated with granular comparative advantage. We analyze whether, as suggested by Gaubert and Itskhoki (2016), the share of the top exporter within an industry is a good predictor of granular comparative advantage. We also investigate whether granular comparative advantage is more likely to emerge when industries command a large fundamental comparative advantage.

We estimate the following regression equation:

$$GCA_{ik} = \alpha ShareTop1_{ik} + \beta FCA_{ik} + \mu_i + \mu_k + \epsilon_{ik} \tag{14}$$

where μ_i and μ_k are region and chapter fixed effects respectively, and ϵ_{ik} is the disturbance term.

Table 4 presents the results of the regression analyses. Column (1) and (2) present the results for 2014, and Columns (3) and (4) for 1998. In column (1) we pool all observations and estimate the regression without region and chapter fixed effects. The top exporter's share is positive and statistically significant; fundamental comparative advantage is positive, although not statistically significant. Column (2) presents the results of estimating (14) with region and chapter fixed effects. The top exporter's share

[24] In Aragon the automobile industry shifted from granular to non-granular from 1998 to 2014. In Asturias organic chemicals becomes an industry with a revealed comparative advantage in 2014. In this industry the granular component is larger than the fundamental component. Finally, in Murcia fuel was a granular sector both in 1998 and 2014. However, the share of fuel in total regional exports was much higher in 2014 than in 1998, raising the share of granular industries in total regional exports.

Table 3: Percentage of granular chapters and share in exports by region, 2014 and 1998

Region	2014			1998		
	% of granular industries	Share of granular in XS >1 industries	% of granular exports	% of granular industries	Share of granular in XS >1 industries	% of granular exports
Average region	37	70	62	37	67	54
Andalusia	33	76	49	31	63	47
Aragon	41	74	35	45	77	82
Asturias	54	94	94	43	71	42
Basque Country	34	57	35	36	60	27
Cantabria	44	76	78	51	88	85
Castile and León	39	60	92	38	71	97
Castile-La Mancha	46	76	81	37	67	82
Catalonia	36	74	29	28	55	14
Extremadura	39	67	72	26	50	79
Galicia	32	68	95	47	91	96
Madrid	26	56	74	34	62	74
Murcia	34	72	71	41	73	18
Navarre	35	56	31	34	66	33
Rioja	45	85	67	40	64	60
Valencia	26	53	25	20	43	10

Source: Authors' estimations using the Customs database. Note: Among industries with export specialization >1, granular industries are defined as those where granular comparative advantage is larger than fundamental comparative advantage.

remains positive and statistically significant, and fundamental comparative advantage remains statistically not significant. The results for 1998 are similar, suggesting that the top exporter's share is a good proxy to identify a granular industry. They also show that fundamental comparative advantage is not positively correlated with granular comparative advantage. Our results are in line with the results obtained by Gaubert and Itskhoki (2016) for French industries.

Table 4: Covariates of granular comparative advantage

	(1)	(2)	(3)	(4)
Top exporter's share	1.745***	1.509***	1.571***	1.370***
	(0.215)	(0.280)	(0.206)	(0.322)
Fundamental comparative advantage	0.131	0.023	0.009	-0.065
	(0.095)	(0.098)	(0.079)	(0.090)
Chapter and region FE	No	Yes	No	Yes
Year	2014	2014	1998	1998
N.observ	581	581	511	511
R squared	0.142	0.365	0.168	0.413

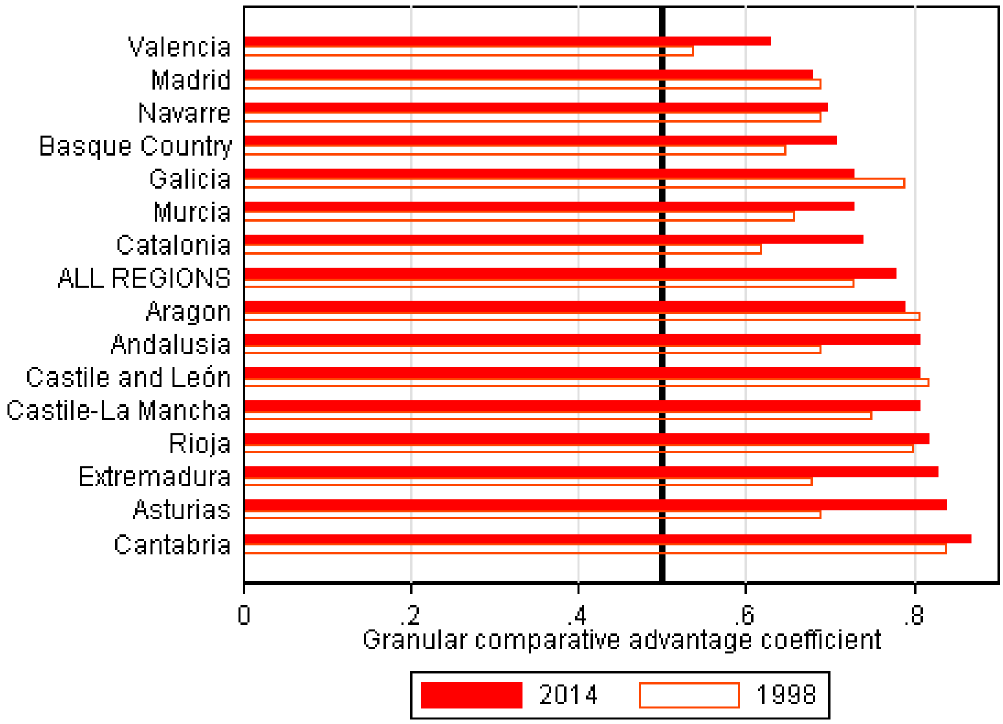
Note: ***, **, * statistically significant at 1%. 5% and 10% respectively. Robust standard errors in parentheses.

Finally, Figure 3 presents the contribution of granular comparative advantage to differences in export specialization across regions. We carry out our baseline analysis using data for 2014. First, we perform the decomposition by pooling all observations; next, we carry out region-specific decompositions. When we pool all observations, 78% of variations in export specialization across chapters and regions are explained by granular comparative advantage and 28% by fundamental comparative advantage. These results suggest that differences in export specialization across regions are mostly explained by granular comparative advantage. The contribution of granular comparative advantage is larger across regions within a country than across countries: 0.78 vs. 0.60. This result confirms the larger role that granularity plays in export specialization at the regional level.

Next, we estimate the contribution of fundamental and granular comparative advantage to the variation in export specialization within each region. We estimate a separate regression for each region. The highest contribution of granular comparative advantage occurs in Cantabria (87%) and the lowest in Valencia (63%). In the large exporting regions, Catalonia and Madrid, the contributions of granular comparative advantage are 74% and 68% respectively. We also carry out the decomposition for 1998. The contribution of granular comparative advantage was slightly lower in 1998 than in 2014.

To end this section, we carry out three robustness analyses. First, we investigate whether the results for Spanish regions are robust to selecting an alternative export destination. Instead of France, we select the second most important destination for Spanish

Figure 3: Contribution of granular comparative advantage to the variation in export specialization, 1998-2014. Regression-based decomposition



Source: Authors' estimations using the Customs database. Note: To calculate the contribution of granular comparative advantage we regress granular comparative advantage on relative exports. To calculate the contribution of fundamental comparative advantage we regress fundamental comparative advantage on relative exports. ALL REGIONS' regression pools the observations from all regions and industries.

exports: Germany. Table A4 in the Appendix presents the results for the percentage of granular chapters, their share in industries with a revealed comparative advantage, and their share in total exports. On average, the results are very similar to those found for France. At the regional level, there are some differences in the share of granular chapters in total exports. Figure A2 presents the results for the contribution of granular comparative advantage to differences in export specialization across regions and industries. The results are very similar to those obtained for France.

Second, we analyze whether the differences between country-level and regional-level results are driven by differences in the characteristics of the sample. As in the country-level analyses, we build a regional sample using average data for the 2008-2013 period, selecting all EU countries as the destination of exports, and using a 21 industry disaggregation. The results are presented in Table A5 and Figure A3 in the Appendix. There is an increase in the percentage of granular industries, in their share in industries with a revealed comparative advantage, in their share in total exports, and in the contribution of granular comparative advantage to differences in export specialization. These results confirm that the role of granular comparative advantage is larger across regions within a country than across countries.

Finally, we investigate whether our results are altered if we use a much more disaggregate industry classification. Instead of the HS 2-digit disaggregation used in the baseline analyses, which distinguishes 96 products, we perform the empirical analyses with the HS 4-digit disaggregation, which distinguishes 1,145 products. As shown in Table A6 in the Appendix, the share of granular industries only rises slightly from 37% (baseline) to 43% in 2014. The contribution of granular comparative advantage to explaining the differences in export specialization also rises slightly from 78% (baseline) to 81% (Figure A4 in the Appendix). These analyses show that our results are robust to a higher product disaggregation level.

4 Conclusions

In this paper we posit a methodology to assess the contribution of Ricardian sectoral productivity differences and firm idiosyncrasies to the definition of export specialization. We show that in a Melitz-type model, where firm productivity is Pareto distributed, and where there is a continuum of firms, export specialization is determined by the ratio of exporter ratios and the ratio of export cost ratios. First, we remove the ratio of export cost ratios by selecting a country and a reference country that have the same relative export costs. Next, we show that the ratio of exporter ratios that we observe in a finite number of firms scenario provides an accurate estimate of the ratio of exporter ratios we would have observed if there had been a continuum of firms. This allow us to estimate fundamental comparative advantage from the ratio of exporter ratios that we observe in the data. Granular comparative advantage, which captures firm-level idiosyncrasies, is computed as the difference between actual export specialization and fundamental comparative advantage.

At the country level, we analyze EU countries' export specialization relative to an average EU country. We find that granular comparative advantage defines export specialization in 29% of industries. We also show that 60% of the differences in export specialization across EU countries are explained by granular comparative advantage. The role of granularity rises when we analyze Spanish regions' export specialization relative to an average Spanish region. This result is in line with the expectation that differences in the sources of fundamental comparative advantage across regions within a country are lower than across countries. Most granular industries remain granular over time, and the contribution of granular comparative advantage to explaining differences in export specialization does not diminish over time. We also find that granular comparative advantage is positively correlated with the share of the top firm in an industry's exports; and granular comparative advantage is not correlated with fundamental comparative advantage.

Our results highlight that countries' and regions' export specialization is not determined solely by variables, such as average productivity or endowments, that might change slowly over time, but also by outstanding firms. They suggest that countries and regions that seek to alter their export specialization should aim to create an environment for new firms to emerge or to attract outstanding firms from other regions or countries.

Appendix A The intensive and the extensive margin of exports in a Melitz-type model where firms' productivity is Pareto distributed

In this Appendix we derive the intensive and extensive margin of exports from a Melitz-type model (Melitz, 2003), where firms' productivity is Pareto distributed.

We assume that firms produce horizontally-differentiated varieties within an industry. Labor is the only production factor. For industry k, the preferences of a representative consumer are given by a constant elasticity of substitution (CES) utility function

$$U = \left(\int_{v \varepsilon \Omega_k} q^{\frac{\sigma-1}{\sigma}} dv \right)^{\frac{\sigma}{\sigma-1}}, \sigma > 1 \tag{A1}$$

where v is a variety that belongs to the set of varieties of industry k (Ω_k), q is the quantity consumed of variety v, and σ is the elasticity of substitution across varieties. It is assumed that σ is common across industries.

In a CES utility framework, the demand of country j of an industry k variety produced in country i is determined by the following expression

$$q_{ijk} = \beta_{jk} Y_j (P_{jk})^{\sigma-1} (p_{ijk})^{-\sigma} \tag{A2}$$

where β_{jk} is the share of income (Y_j) that country j devotes to industry k, and P_{jk} is the price index of industry k varieties in country j.

Since there is monopolistic competition, firms set prices as a constant mark-up over marginal costs

$$p_{ijk} = \frac{\sigma}{\sigma - 1} c_{ijk} \tag{A3}$$

where c_{ijk} is the marginal cost of selling a unit of an industry k variety in country j. This cost is determined by the following expression:

$$c_{ijk} = \frac{\tau_{ijk} w_i}{z} \tag{A4}$$

where τ_{ijk} is an iceberg-type trade cost, denoting the units of an industry k variety that should be sent from country i to country j to ensure that one unit arrives; w_i is the wage in country i, and z is the productivity of the firm.

Combining (A2), (A3) and (A4), exports per firm are given by

$$x_{ijk} = \beta_{jk} Y_j (P_{jk})^{\sigma-1} \Big(\frac{\sigma}{\sigma-1} \frac{\tau_{ijk} w_i}{z} \Big)^{-\sigma} \tag{A5}$$

Firms will export to country j if they obtain profits from export. This happens when the following condition is met

$$(p_{ijk} - c_{ijk}) q_{ijk} > F_{ijk} \tag{A6}$$

where F_{ijk} is the fixed costs that a firm in i has to cover if it wants to export a k industry variety to country j. Substituting (A2), (A3) and (A4) in (A6) we get,

$$z > \Big(\frac{F_{ijk}}{\mu \beta_j Y_j} \Big)^{(1/\sigma-1)} \Big(\frac{w_i \tau_{ijk}}{P_{jk}} \Big) \tag{A7}$$

where $\mu = (\sigma - 1)^{\sigma-1} \sigma^{-\sigma}$.

If we substitute the inequality in (A7) with an equality, we get the threshold productivity that firms in country i should reach to export a variety of industry k to country j

$$\underline{z}_{ijk} = \Big(\frac{F_{ijk}}{\mu \beta_{jk} Y_j} \Big)^{(1/\sigma-1)} \Big(\frac{w_i \tau_{ijk}}{P_{jk}} \Big) \tag{A8}$$

Following Chaney (2008), we assume that there is a large exogenous pool of firms, M_i, that can potentially enter any industry in country i. The value of industry k exports from country i to country j will be determined by the sum of exports of the potential entrants that reach a productivity equal or above the threshold productivity to export.

If productivity is Pareto distributed, the probability density function is given by:

$$G(z) = \frac{\theta \varphi_{ik}^{\theta}}{z^{\theta+1}} \tag{A9}$$

where φ_{ik} is the minimum productivity that firms in country i can get in industry k, and θ measures the heterogeneity in the distribution of productivity.[25]

The amount of k industry exports from region i to country j is determined by:

$$X_{ijk} = M_i \int_{\underline{z}_{ijk}}^{-\infty} \beta_{jk} Y_j (P_{jk})^{\sigma-1} \Big(\frac{\sigma}{\sigma-1} \frac{\tau_{ijk} w_i}{z} \Big)^{-\sigma} \frac{\theta (\varphi_{ik})^{\theta}}{z^{\theta+1}} dz \tag{A10}$$

Solving the integral in (A10), we get

[25]For stability, it is also assumed that $\theta > \sigma - 1$.

$$X_{ijk} = M_i \beta_{jk} Y_j \frac{\theta}{\theta - \sigma - 1} (\varphi_{ik})^{\theta} \underline{z}_{ijk}^{\sigma-\theta-1} \Big(\frac{\sigma}{\sigma - 1} \frac{P_{jk}}{\tau_{ijk} w_i} \Big)^{\sigma-1} \tag{A11}$$

With a Pareto distribution, the number of firms in country i that export an industry k variety to country j is given by:

$$N_{ijk} = M_i \int_{\underline{z}_{ijk}}^{-\infty} \frac{\theta(\varphi_{ik})^{\theta}}{z^{\theta+1}} dz = M_i \Big(\frac{\underline{z}_{ik}}{\varphi_{ik}} \Big)^{-\theta} \tag{A12}$$

To get the average exports per firm, also denoted as the intensive margin of exports, we divide (A11) by (A12)

$$\overline{x}_{ijk} = \frac{X_{ijk}}{N_{ijk}} = \beta_{jk} Y_j \frac{\theta}{\theta - \sigma - 1} \underline{z}_{ijk}^{\sigma-1} \Big(\frac{\sigma}{\sigma - 1} \frac{P_{jk}}{\tau_{ijk} w_i} \Big)^{\sigma-1} \tag{A13}$$

If we substitute (A8) in (A13), the variables β_{jk}, Y_j, P_{jk}, τ_{ijk} and w_i cancel out, leaving the expression

$$\overline{x}_{ijk} = \Big(\frac{\theta\sigma}{(\theta - \sigma + 1)} \Big) F_{ijk} \tag{A14}$$

Appendix B Coefficient of variation of the ratio of exporter ratios

We estimate fundamental comparative advantage using the ratio of exporter ratios:

$$\frac{N_{ijk}/N_{ijk'}}{N_{i'jk}/N_{i'jk'}} \tag{A15}$$

To facilitate the analysis, we express this ratio as follows:

$$\frac{N_{ijk}/N_{ijk'}}{N_{i'jk}/N_{i'jk'}} = (N_{ijk} N_{i'jk'}) \Big(\frac{1}{N_{ijk'} N_{i'jk}} \Big) = (a_1 a_2) \Big(\frac{1}{b_1 b_2} \Big) = AB \tag{A16}$$

where $A = a_1 a_2$ and $B = \frac{1}{b_1 b_2}$

To express the coefficient of variation of [AB] we need the expectation of [AB] and the variance of [AB].

If A and B are independent, the expectation of (A15) is:

$$E[AB] = E[A]E[B] = E[a_1 a_2] E\Big[\frac{1}{b_1 b_2}\Big] = E[a_1]E[a_2]E\Big[\frac{1}{b_1}\Big]E\Big[\frac{1}{b_2}\Big] \tag{A17}$$

And the variance,

Table A1: Numerical simulations of the ratio of exporter ratios coefficient of variation

Simulation	M_i	$M_{i'}$	θ	$\underline{z}_{ijk}/\varphi_{ik}$	$\underline{z}_{ijk'}/\varphi_{ik'}$	$\underline{z}_{i'jk}/\varphi_{i'k}$	$\underline{z}_{i'jk'}/\varphi_{i'k'}$	Coefficient of variation (mean)
1	100	100	5	1.11	1.11	1.11	1.11	0.268
2	1,000	1,000	5	1.11	1.11	1.11	1.11	0.087
3	100	1,000	5	1.11	1.11	1.11	1.11	0.197
4	100	100	4	1.11	1.11	1.11	1.11	0.253
5	100	100	5	1.50	1.50	1.50	1.50	0.668
6	100	100	5	1.50	1.11	1.50	1.11	0.496

Note: In each simulation, we calculate the variation coefficient over a sample of 1,000 random numbers generated by the Poisson distribution. The coefficient of variation presented in the table is the average of 1,000 simulations.

$$Var[AB] = E[A^2]E[B^2] - (E[A])^2(E[B])^2 = E[(a_1a_2)^2]E[\Big(\frac{1}{b_1b_2}\Big)^2] - (E[a_1])^2(E[a_2])^2\Big(E\Big[\frac{1}{b_1}\Big]\Big)^2\Big(E\Big[\frac{1}{b_2}\Big]\Big)^2 \tag{A18}$$

The coefficient of variation can be expressed as:

$$cv[AB] = \frac{\sqrt{E[(a_1a_2)^2]E\Big[\Big(\frac{1}{b_1b_2}\Big)^2\Big] - (E[a_1])^2(E[a_2])^2\Big(E\Big[\frac{1}{b_1}\Big]\Big)^2\Big(E\Big[\frac{1}{b_2}\Big]\Big)^2}}{E[a_1]E[a_2]E\Big[\frac{1}{b_1}\Big]E\Big[\frac{1}{b_2}\Big]} \tag{A19}$$

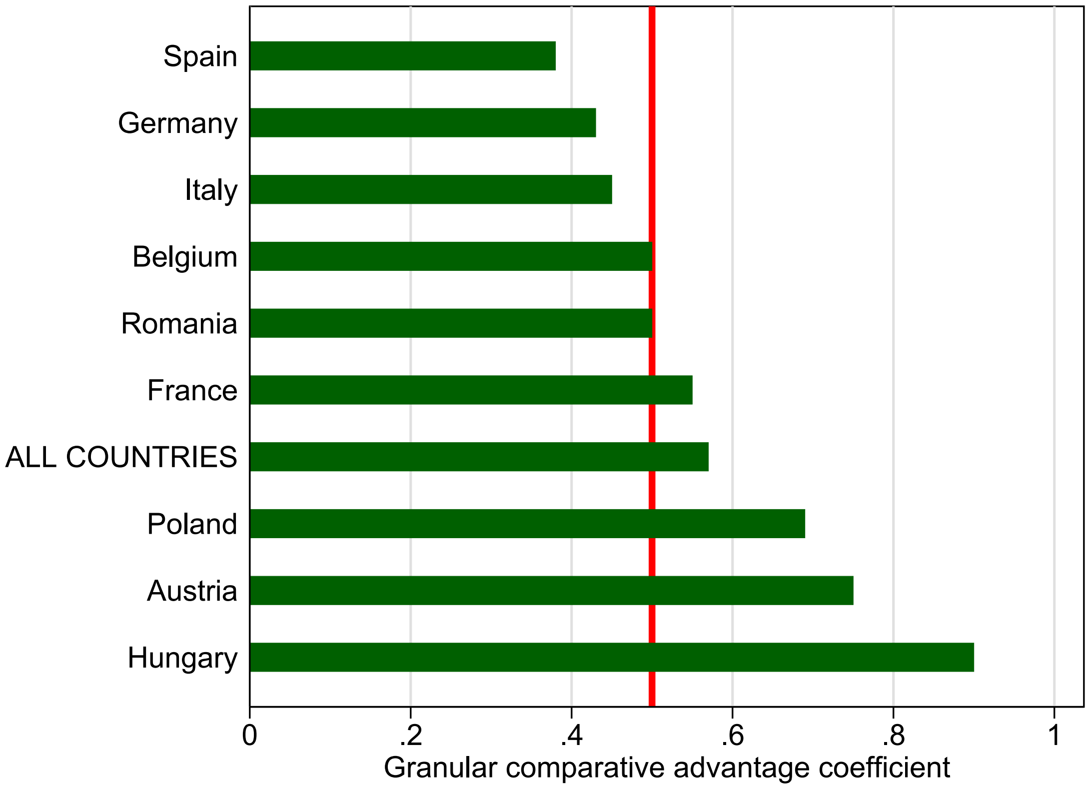
To gauge the range of values the coefficient of variation may take in (A19), we use numerical simulations with random numbers generated from a Poisson distribution using different λ parameters. For each λ parameter, we calculate the variation coefficient over a sample of 1,000 random numbers. Then, we repeat the simulation 1,000 times and calculate an average variation coefficient. Table A1 presents the results of these simulations.

In Simulation 1 the number of draws is very low (M_i=100). We use the baseline shape parameter, $\theta = 5$, and the baseline $\underline{z}_{ijk}/\varphi_{ik}$ value, 1.11, which is common to both industries. We assume that the number of draws within a country is the same for analyzed industry k and reference industry k'. In addition, analyzed country i and reference country i' have the same number of draws ($M_i = M_{i'}$). Simulation 1 yields a 0.268 variation coefficient. Since distributions with a variation coefficient less than one are considered to be low-variance, we can qualify this value as very low. Simulation 2 raises the number of draws to 1,000 in both countries. With a less restrictive number of draws, the variation coefficient drops to 0.087. Simulation 3 combines a restrictive number of draws in i and a less restrictive number of draws in i'. The coefficient of variation remains very low. Simulation 4 analyzes whether the results are sensitive to shape parameter θ. We reduce the value of the parameter to 4. This yields a coefficient of variation that is slightly lower than the one obtained in Simulation 1. In Simulation 5 we raise the threshold productivity to export/fundamental productivity ratio to 50%. The variation coefficient rises to 0.668, but is still below the benchmark value of 1. When we combine a higher threshold/fundamental productivity ratio in one industry with a lower ratio in the other, the coefficient of variation drops to 0.496.

These simulations show that even if we consider scenarios with a low number of draws and large differences between threshold and fundamental productivities, the coefficient of variation remains small.

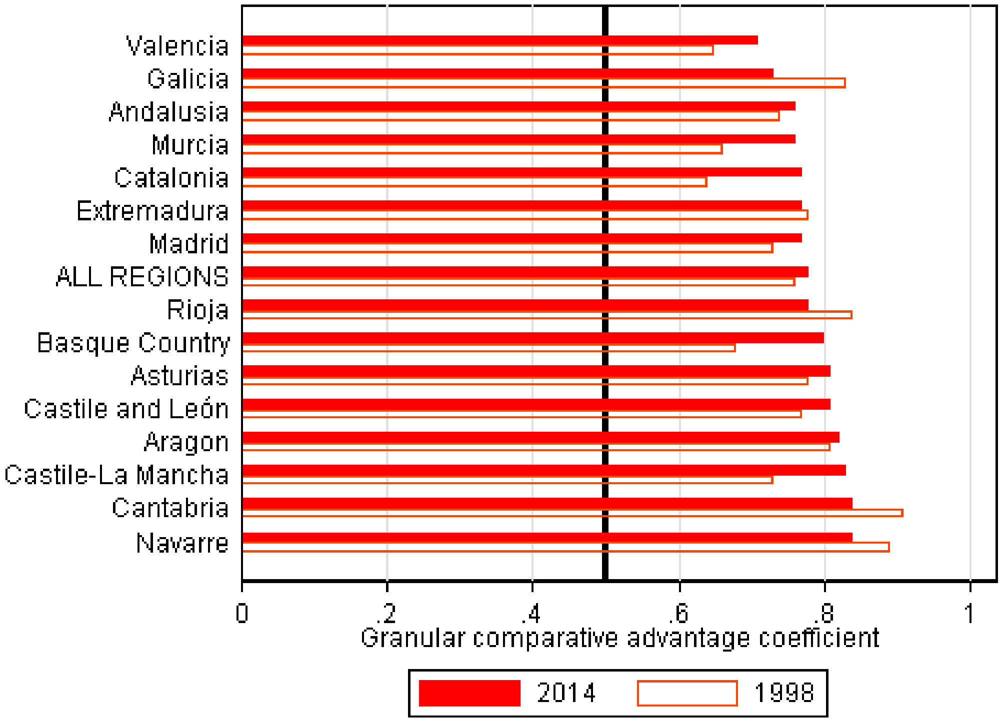
Map A1: NUTS II Regions of Spain

Figure A1: Robustness. Small firms removed. Contribution of granular comparative advantage to variation in export specialization. Regression-based decomposition (Average 2008-2013)



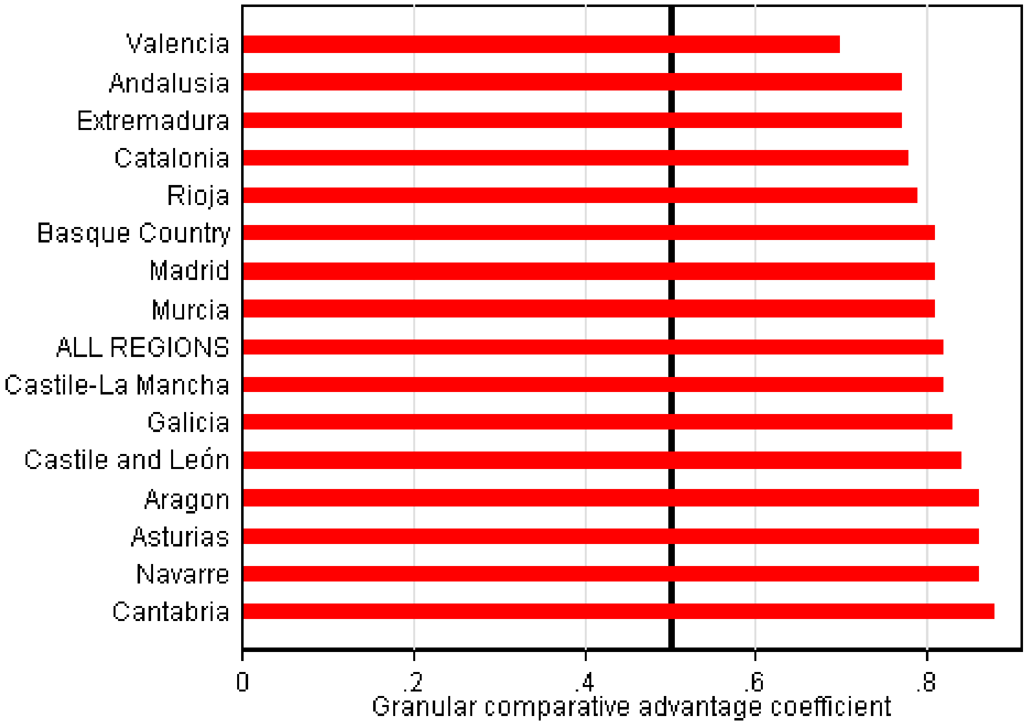
Source: Authors' estimations using the OECD-Eurostat Trade by Enterprise Characteristics database. Note: To calculate the contribution of granular comparative advantage we regress granular comparative advantage on export specialization. ALL COUNTRIES' regression pools the observations from all countries and industries.

Figure A2: Robustness. Germany as destination country. Contribution of granular and fundamental comparative advantage to the variation in export specialization, 1998-2014. Regression-based decomposition



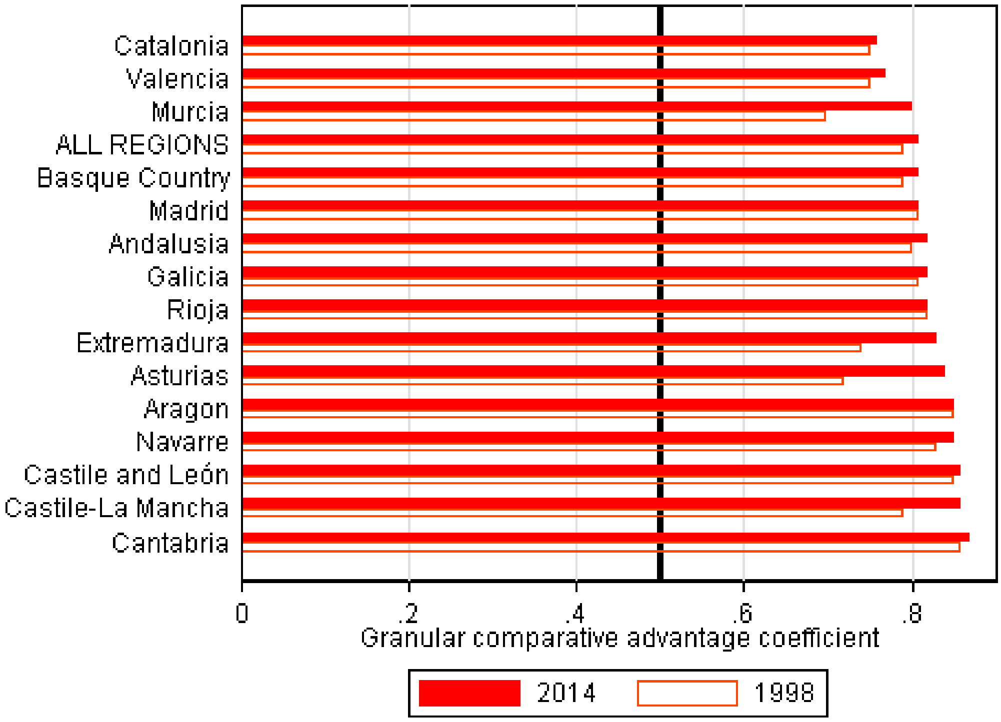
Source: Authors' estimations using the Customs database. Note: To calculate the contribution of granular comparative advantage we regress granular comparative advantage on relative exports. ALL REGIONS' regression pools the observations from all regions and industries.

Figure A3: Robustness. Regional sample has the characteristics of the country sample. Contribution of granular and fundamental comparative advantage to variation in export specialization. Regression-based decomposition (Average 2008-2013; intra-EU exports)



Source: Authors' estimations using the Customs database. Note: To calculate the contribution of granular comparative advantage we regress granular comparative advantage on relative exports. ALL REGIONS' regression pools the observations from all regions and industries.

Figure A4: Robustness. HS 4-digit industry disaggregation. Contribution of granular and fundamental comparative advantage to variation in export specialization. Regression-based decomposition, 1998-2014



Source: Authors' estimations using the Customs database. Note: To calculate the contribution of granular comparative advantage we regress granular comparative advantage on relative exports. ALL REGIONS' regression pools the observations from all regions and industries.

Table A2: OECD-Eurostat Database: Summary statistics for the EU countries in the sample

Country	Period	Number of manufacturing exporters to the EU	Intra-EU manufacturing exports (million USD)	Share of export value to the EU by road
Austria	2008-2013	8211	75367	n.a.
Belgium	2011-2013	11719	123111	81
France	2008-2013	17248	207773	n.a.
Germany	2008-2013	51924	427052	85
Hungary	2008-2013	8900	53672	93
Italy	2008-2013	72012	213979	88
Lithuania	2008-2013	2521	11426	65
Netherlands	2009-2013	10741	105570	n.a.
Poland	2008-2013	22247	93386	88
Portugal	2008-2013	12534	29498	78
Romania	2008-2012	7376	27550	89
Slovakia	2008-2013	3064	31649	80
Slovenia	2009-2013	n.a.	n.a.	n.a.
Spain	2008-2013	18459	116616	76

Source: OECD-Eurostat Trade by Enterprise Characteristics database and Eurostat's Trade database. Note: Number of exporters and intra-EU exports are the average for the period. Share of exports value by road corresponds to 2013. n.a.: Not available. Some countries, due to confidentiality problems in some industries, do not provide aggregate data either.

Table A3: Robustness. Small firms removed. Granular industries by country (Average 2008-2013)

Country	% of granular industries	Share of granular in XS >1 industries	% of granular exports
Average country	31	63	49
Austria	33	78	38
Belgium	29	60	45
France	14	33	22
Germany	29	50	53
Hungary	38	80	71
Italy	33	58	24
Poland	33	70	65
Romania	33	70	57
Spain	33	58	53

Source: Authors' calculations using the OECD-Eurostat Trade by Enterprise Characteristics database. Note: XS=Export specialization. Among industries with export specialization >1, granular industries are defined as those where granular comparative advantage is larger than fundamental comparative advantage.

Table A4: Robustness. Germany as destination country. Granular industries by region, 2014 and 1998

Region	2014			1998		
	% of granular industries	Share of granular in XS >1 industries	% of granular exports	% of granular industries	Share of granular in XS >1 industries	% of granular exports
Average region	39	73	68	37	67	51
Andalusia	34	69	41	37	69	33
Aragon	42	84	96	37	73	91
Asturias	42	83	79	45	68	77
Basque Country	42	70	22	35	66	15
Cantabria	43	86	93	42	72	90
Castile and León	40	71	92	33	60	61
Castile-La Mancha	38	73	53	42	69	74
Catalonia	43	74	59	30	55	15
Extremadura	32	61	85	44	76	88
Galicia	35	71	72	48	87	85
Madrid	36	66	60	34	66	27
Murcia	43	76	56	31	53	10
Navarre	38	69	90	43	74	39
Rioja	32	70	56	35	69	60
Valencia	34	70	77	32	58	46

Source: Authors' estimations using the Customs database. Note: Among industries with export specialization >1, granular industries are defined as those where granular comparative advantage is larger than fundamental comparative advantage.

Table A5: Robustness. Regional sample has the characteristics of the country sample. Granular chapters by region (average 2008-2013; Intra-EU exports)

Region	% of granular industries	Share of granular in XS >1 industries	% of granular exports
Average region	42	74	68
Andalusia	49	68	63
Aragon	51	78	92
Asturias	71	100	99
Basque Country	42	74	79
Cantabria	64	90	97
Castile and León	30	61	78
Castile-La Mancha	11	56	18
Catalonia	41	78	21
Extremadura	48	72	85
Galicia	47	80	81
Madrid	60	86	88
Murcia	40	82	64
Navarre	33	69	84
Rioja	33	82	74
Valencia	11	30	4

Source: Authors' calculations using the Customs database. Note: Granular industries are defined as those where granular comparative advantage is larger than fundamental comparative advantage.

Table A6: Robustness. HS 4-digit product disaggregation. Granular industries by region, 2014 and 1998

Region	2014			1998		
	% of granular industries	Share of granular in XS >1 industries	% of granular exports	% of granular industries	Share of granular in XS >1 industries	% of granular exports
Average region	43	77	84	39	70	71
Andalusia	39	79	67	42	76	61
Aragon	45	79	87	45	72	93
Asturias	50	82	96	34	56	66
Basque Country	42	77	71	38	70	61
Cantabria	43	77	95	54	84	96
Castile and Len	39	73	95	40	69	98
Castile-La Mancha	43	79	91	41	73	86
Catalonia	34	70	52	31	67	46
Extremadura	53	86	98	30	58	89
Galicia	42	78	95	44	83	96
Madrid	39	77	92	39	75	86
Murcia	43	78	86	32	59	33
Navarre	44	80	77	41	74	80
Rioja	43	73	87	38	73	72
Valencia	38	72	64	37	74	47

Source: Authors' estimations using the Customs database. Note: Among industries with export specialization >1, granular industries are defined as those where granular comparative advantage is larger than fundamental comparative advantage.

References

Albornoz, F., Fanelli, S., and Hallak, J. C. (2016). Survival in export markets. *Journal of International Economics*, 102:262 – 281.

Araújo, S. and Gonnard, E. (2011). Selling to foreign markets: A portrait of OECD exporters. *OECD Statistics Brief*, (16):1–8.

Arkolakis, C., Costinot, A., and Rodríguez-Clare, A. (2012). New trade models, same old gains? *The American Economic Review*, 102(1):94–130.

Artal-Tur, A., Castillo-Gimnez, J., Llano-Verduras, C., and Requena-Silvente, F. (2011). The factor content of regional bilateral trade: The role of technology and demand. *International Review of Economics & Finance*, 20(2):157 – 172.

Bas, M., Mayer, T., and Thoenig, M. (2016). From micro to macro: Demand, supply, and heterogeneity in the trade elasticity. *CEPR Discussion Paper No. DP10637.* Center for Economic Policy Research.

Bernard, A. B., Jensen, J. B., Redding, S. J., and Schott, P. K. (2007a). Firms in international trade. *The Journal of Economic Perspectives*, 21(3):195–130.

Bernard, A. B., Redding, S. J., and Schott, P. K. (2007b). Comparative advantage and heterogeneous firms. *The Review of Economic Studies*, 74(1):31.

Cadot, O., Iacovone, L., Pierola, M. D., and Rauch, F. (2013). Success and failure of African exporters. *Journal of Development Economics*, 101:284 – 296.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review*, 98(4):1707–21.

Combes, P.-P. and Lafourcade, M. (2005). Transport costs: measures, determinants, and regional policy implications for France. *Journal of Economic Geography*, 5(3):319.

Coşar, A. K. and Fajgelbaum, P. D. (2016). Internal geography, international trade, and regional specialization. *American Economic Journal: Microeconomics*, 8(1):24–56.

Costinot, A., Donaldson, D., and Komunjer, I. (2012). What goods do countries trade? A quantitative exploration of Ricardo's ideas. *The Review of Economic Studies*, 79(2):581–608.

Courant, P. N. and Deardorff, A. V. (1992). International trade with lumpy countries. *Journal of Political Economy*, 100(1):198–210.

De Lucio, J., Mínguez, R., Minondo, A., and Requena, F. (2017). The granularity of Spanish exports. *Working Papers in Applied Economics 2017-01*. Universitat de València.

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.

Esteve-Pérez, S., Requena-Silvente, F., and Pallardó-López, V. J. (2013). The duration of firm-destination export relationships: Evidence from Spain, 1997-2006. *Economic Inquiry*, 51(1):159–180.

Eurostat (2016). *Compilers guide on European statistics on international trade in goods by enterprise characteristics (TEC)*. Eurostat, Luxembourg.

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

Fernandes, A. M., Klenow, P. J., Meleshchuk, S., Pierola, M. D., and Rodríguez-Clare, A. (2015b). The intensive margin in trade: Moving beyond Pareto.

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

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

Görg, H., Kneller, R., and Muraközy, B. (2012). What makes a successful export? Evidence from firm-product- level data. *Canadian Journal of Economics/Revue canadienne d'conomique*, 45(4):1332–1368.

Guillaume, G., Mirza, D., Turban, S., and Zignano, S. (2008). International transportation costs around the world: a new CIF/FoB rates dataset. *CEPII Working Paper*.

Head, K. and Mayer, T. (2014). Gravity equations: Workhorse, toolkit, and cookbook, ch. 3 in Handbook of International Economics, Gopinath, G, E. Helpman and K. Rogoff (eds), vol. 4, 131-95.

Head, K., Mayer, T., and Thoenig, M. (2014). Welfare and trade without Pareto. *The American Economic Review*, 104(5):310–316.

Helpman, E., Melitz, M., and Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *The Quarterly Journal of Economics*, 123(2):441–487.

Hummels, D. and Lugovskyy, V. (2006). Are matched partner trade statistics a usable measure of transportation costs? *Review of International Economics*, 14(1):69–86.

Jones, R. W. and Neary, J. P. (1984). The positive theory of international trade. In Ronald W., J. and Peter B., K., editors, *Handbook of International Economics*, volume 1, pages 1–62. Elsevier, North-Holland.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Minondo, A. (2017). Fundamental versus granular comparative advantage: An analysis using chess data. *Kyklos*. forthcoming.

Nigai, S. (2017). A tale of two tails: Productivity distribution and the gains from trade. *Journal of International Economics*, 104:44 – 62.

Requena, F., Artal, A., and Castillo, J. (2008). Testing Heckscher-Ohlin-Vanek model using Spanish regional data. *International Regional Science Review*, 31(2):159–184.