# Cities Trade Pattern*

Jorge Díaz-Lanchas† Carlos Llano‡ Asier Minondo§
Francisco Requena¶

September 29, 2015
*Preliminary draft*

**Abstract**

Recent theoretical models predict that large cities should specialize in more skill-intensive goods. In this paper we use international trade data for Spanish urban areas in 2012 to test this hypothesis. We show that trade specialization in skill-intensive goods raises with urban area population. We also find that larger cities tend to export more complex goods, and specialize in products that make intensive use of social and cognitive skills, and combine tasks that should be performed geographically close. These results confirm that city size plays a very important role in determining trade patterns.

**JEL**: F11, F14, R12
**Keywords**: urban areas, exports, Spain, comparative advantage, complexity

*The authors thank Daniel Sánchez Serra for providing the correspondence between Spanish municipalities and urban areas. They also thank Francisco Olarte for preparing the SABI data used in this study and Raúl Mínguez for providing us with international trade data at the urban area level. We also thank the comments received from participants at the XVIII Encuentro de Economía Aplicada. They gratefully acknowledge financial support from the Spanish Ministry of Economy and Competitiveness (MINECO ECO2013-46980-P, co-financed with FEDER), the Basque Government Department of Education, Language policy and Culture (IT629-13) and the FPU-2010 Scholarship for pre-doctoral Education Training by the Spanish Ministry of Education.

†Díaz-Lanchas: Department of Economics. Universidad Autónoma de Madrid. 28049 Cantoblanco. Madrid. Email:jorge.diaz@uam.es.

‡Llano: Department of Economics. Universidad Autónoma de Madrid. 28049 Cantoblanco. Madrid. Email: carlos.llano@uam.es.

§Minondo: Corresponding author. Deusto Business School, University of Deusto, Camino de Mundaiz 50, 20012 Donostia - San Sebastián, Spain. Email: aminondo@deusto.es.

¶Requena: The University of Sheffield, Western Bank, Sheffield, S10 2TN, UK. Email: f.requena@sheffield.ac.uk

# 1 Introduction

Cities are one of the most salient examples of agglomeration economies. Despite the reduction in communication costs, the value of proximity does not seem to decline and an increasing share of people live in cities (Gaspar and Glaeser, 1998; Glaeser, 2011). However, there are substantial differences across cities. In particular, the literature shows that larger cities have a higher share of skilled workers, host more productive firms and pay higher wages than smaller cities (Glaeser and Resseger, 2010; Behrens and Robert-Nicoud, 2015).

This paper analyzes whether cities also differ in trade patterns. As explained by Davis and Dingel (2014), if the productivity of skilled workers raises with the skill intensity of the industry, and if they are more productive when surrounded by other skilled workers, large cities will have a comparative advantage in skill-intensive goods. As comparative advantage governs trade patterns in neoclassical trade models, differences in comparative advantage should be reflected in cities export specialization. Moreover, export specialization seems a better indicator to show the differences in comparative advantage, than production or employment indicators, because cities are expected to export the goods in which they are more productive (Hausmann et al., 2007).

This paper uses international trade data of Spanish urban areas for the year 2012 to test whether larger cities export more skill-intensive goods than smaller cities. We perform empirical tests on urban areas exports by industry, and urban areas exports by industry and destination. These tests confirm that large cities export more skill-intensive goods. Following the insights of previous studies, we also analyze cities trade pattern from additional angles. First, we show that large cities export goods that are intensive in the use of social and cognitive skills (Bacolod et al., 2009). Second, large cities also export goods that combine tasks that should be performed geographically close (Kok and Weel, 2014). Third, we also show that large cities export goods that require the combination of a large number of different skills (Jacobs, 1969; Duranton and Puga, 2001; Minondo and Requena-Silvente, 2013; Hausmann and Hidalgo, 2014).

This paper contributes to both the urban and the trade literature. Regarding the urban literature, as far as we know, we are the first paper to show that there is heterogeneity across cities regarding trade patterns. Our findings confirm the conclusions of Davis and Dingel (2014) which argue that large cities should have a comparative advantage in skill-intensive goods. Our results also give support to validate previous findings by Bacolod et al. (2009) and Kok and Weel (2014) that argued that cities should specialize in industries that make intensive use of social and cognitive skills, and tasks that should be performed in proximity. The paper also contributes to the trade literature, showing that skilled workers incentives to agglomerate might lead to differences in trade patterns within countries. Our paper also provides another example of complementarity, or log-supermodularity, at the country level, to explain differences in trade patterns (Costinot, 2009a). It also confirms that complexity explains trade pattern differences, not only across countries, as in Costinot (2009b) or Minondo and Requena-Silvente (2013), but also within countries.

The remainder of the paper is organized as follows. Section 2 presents the backbone of the model that drives the empirical tests. Section 3 describes the different data sources used in the empirical analyses. Section 4 reports the main findings and the results of robustness analyses. The last section concludes.

# 2 Theoretical framework

We follow the theoretical framework developed in Davis and Dingel (2014). The model has two main features. First, the productivity advantage of skilled workers relative to non-skilled workers raises with the skill intensity of the industry. That is, production is log-supermodular in workers skills and industries skill intensity. Second, skilled workers have an incentive to concentrate, because they are more productive when they work close to other skilled workers.[1] Combining these elements, the model yields an equilibrium characterized by a positive relationship between city population and the share of workers employed in skill-intensive industries. In the Davis-Dingel model there are no trade costs and preferences are homothetic. Hence, if free trade is allowed, the model predicts a positive relationship between city size and skill-intensive exports.

Davis and Dingel (2014) derive two empirical tests to assess the validity of the model. The first test, named the "elasticity test", analyzes whether the population elasticity of exports is increasing in industries' skill intensity. If we compare a skill-intensive industry with an unskilled-intensive industry, we expect exports to rise more with population in the former than in the latter. In the second test, named the "pairwise comparison test", cities are sorted according to population and divided into bins. If we select two bins and two industries, the model predicts that the ratio of skill/unskilled-intensive exports should be equal or higher in the large-population bin than in the low-population bin.

These empirical tests use information about industries' skill intensities and urban areas total exports by industry. To take advantage of our export data, which is disaggregated by urban areas, industries and destinations, we turn to international trade models that have incorporated the insights from the mathematics of complementarity, or log-supermodularity (Costinot, 2009a). According to this literature, countries specialize in industries that make intensive use of a characteristic or factor in which the country is

[1]This larger productivity might be due to the higher opportunities of learning and experimentation (Glaeser, 1999; Duranton and Puga, 2001; Davis and Dingel, 2012).

well endowed. For example, Costinot (2009b) shows that countries with a higher institutional quality and human capital tend to specialize in more complex goods, defined as those that combine a larger number of tasks. He derives an econometric equation that allows to verify the complementarity between industries and exporters characteristics. We adapt that equation to our analysis to test the relationship between cities size and specialization in skill-intensive industries. The regression equation is defined as

$$\ln x_{ij}^{k} = \beta pop_i s^k + \mu_{ij} + \mu_{jk} + \epsilon_{ij}^{k} \tag{1}$$

where $lnx_{ij}^{k}$ are industry $k$ (log) exports from city $i$ to country $j$, $pop_i$ is the population of city $i$, $s^k$ is the skill intensity of industry $k$, $\mu_{ij}$ is an origin city-country of destination fixed-effect, $\mu_{jk}$ is a country of destination-industry fixed effect, and $\epsilon_{ij}^{k}$ the error term. Exports log-supermodularity in population and skill intensity is captured by the interaction term $pop_i s^k$, which measures the positive relationship between city size and skill-intensive exports. The exporter-importer fixed effect, $\mu_{ij}$, captures the wages in city $i$ and the trade barriers between city $i$ and importer $j$. The importer-industry fixed effect captures the differences in preferences and barriers across importing countries and industries.

In addition to skill intensity, which is the main variable which distinguishes industries in our analysis, we also study other industry-level variables that might describe cities trade pattern. First, Bacolod et al. (2009) argue that larger cities reward cognitive and social skills above other skills, such as physical strength. As explained before, a reason that might lead skilled workers to concentrate geographically is the opportunity to learn from others. As argued by Bacolod et al. (2009), learning opportunities will be larger for workers with higher cognitive skills. As learning also happens discussing ideas with others, having social or soft skills might enhance the scope and deepness of social relationships. Hence, cities might particularly attract workers that command cognitive and social skills, providing them a comparative advantage in industries that make intensive use of these skills. Second, as argued in the classical work of Jacobs (1969), large cities also command a more diversified range of skills. This diversity might render large cities a comparative advantage in activities that are in their experimental stage, as in Duranton and Puga (2001), or in industries that demand the combination of different tasks, as in Costinot (2009a), Minondo and Requena-Silvente (2013) and Hausmann and Hidalgo (2014). Finally, Kok and Weel (2014) argue that some tasks should be performed at a close geographical distance. As cities concentrate a large number of skilled workers in a small geographical area, they might have an advantage in activities that make intensive use of tasks that should be performed in proximity.

# 3 Data

In this section we present the classification of urban areas and the data on international trade and industry characteristics that we use in the empirical analyses. At the end of this section we present some figures to motivate the empirical tests.

*Urban areas*

We use the functional urban areas identified by the OECD for Spain. The OECD follows a three step approach to define functional urban areas (OECD, 2012). First, they identify densely populated municipalities. Second, they aggregate densely populated municipalities into an urban area if more than 15% of the population of one municipality commutes to work in the other municipality. Finally, municipalities that have a low population density are assigned to an urban area if at least 15% of their employed population work in that urban area.

Figure 1 shows the 76 urban areas identified by the OECD in Spain's map.[2] They account for 67% of the Spanish population in 2012. Spain has 2 large metropolitan areas: Barcelona and Madrid (with a population of 1.5 million of more), 6 metropolitan areas (with a population between 500,000 and 1.5 million), 22 medium-size urban areas (with a population between 200,000 and 500,000), and 46 small urban areas (with a population below 200,000 people). All the large metropolitan and metropolitan areas, and most of the medium-size urban areas, are located around a province capital. Data on urban area population and education levels are obtained from the Spanish Statistical Institute (INE) Census.

To test the robustness of our empirical results, we also use the (large) urban areas identified by the Spanish Ministry of Public Works. They are composed by 86 (large) urban areas which accounted for 68.5% of the Spanish population in 2012.[3]

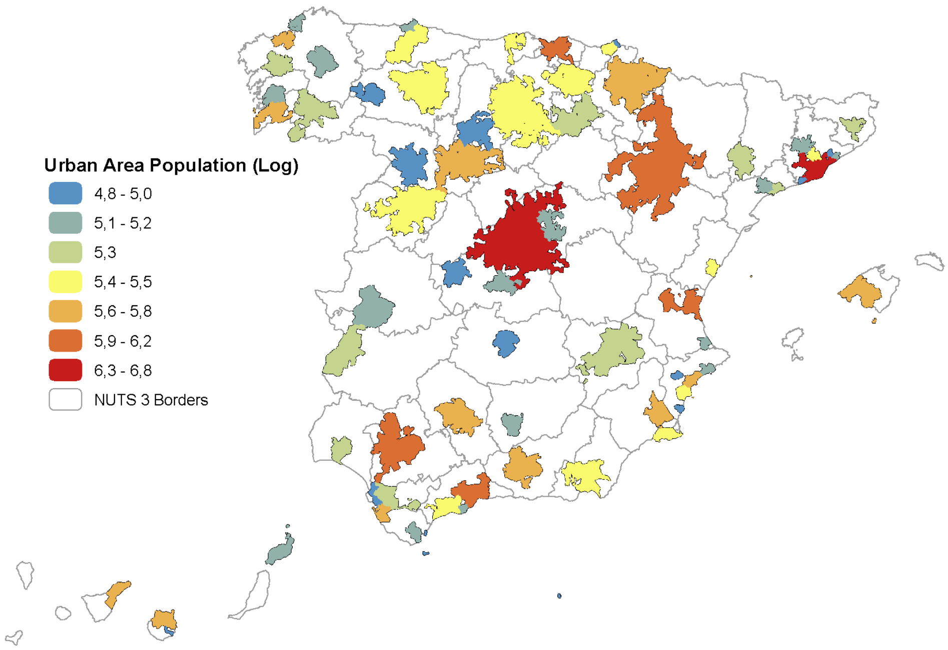
*Exports*

We calculate exports at the urban area level using firm-level data included in the Directory of the Spanish Exporting and Importing firms, which is maintained by the Spanish Chamber and the Inland Revenue Agency. Firms included in the Directory report the annual value of exports per product (at the 8-digit level of the European Common Nomenclature classification) and destination. The database also reports the location of the firm, so we can add-up exports by product and destination at the urban

[2] Table A1 in the appendix presents the list of urban areas and their population in 2012.

[3] Large urban areas should have a municipality with at least 50,000 inhabitants. Other municipalities are added to the urban area if they are geographically close to the main municipality and have at least 1,000 inhabitants.

**Figure 1:** Urban areas in Spain (OECD classification)




Source: Own elaboration from OECD Urban Areas Classification.

area level. The export transactions that are can be allocated to an urban area represent 73% of all export transactions included in the Directory, and they represent 81% of the total export value. The Directory includes exports data for 5,896 firms, which represent approximately 7% of Spanish exporters in 2011.[4] These firms account for 27% of Spanish exports in 2012. The largest effort to include firms in the Directory took place during the years 2003 and 2004. In that time, almost 30,000 exporters were approached by mail by the Inland Revenue Agency and the Chamber of Spain asking them permission to include their trade data in the Directory. In these large group of firms they included the most important exporters, which explains the highest representativeness of the Directory in terms of value than in terms of exporters. The large number of firms that were approach to participate in the Directory ensures that the sample does not have any serious bias in the representativeness at the urban area/industry combination.

Our database has some limitations that should be highlighted. First, there are three chapters of the HS 2-digit classification where exports reported by the Directory are larger than total Spanish exports reported by the Inland Revenue Agency. These chapters are silk (HS 50), tin (HS 80) and cermets (HS 81). This difference might be due to a

[4] According to the Inland Revenue Agency there were 137,528 export operators in 2012. However, in the year 2011 the Inland Revenue Agency crossed the operators (123,128) with the firms included in the Spanish Central Firms' Directory, and found that only in 83,725 were firms. Hence, we use this figure as our benchmark.

misclassification of exports by the firms included in the Directory. We decide to remove these chapters from the database. Second, the Directory reports the postal code of the municipality where the headquarter of the firm is located. It might be the case that the headquarter of the firm is located in a urban area, but the factory where goods are manufactured and exported are in another urban area. In these cases, exports might be allocated to urban areas in erroneously. This problem might be specially problematic for the urban areas of Madrid and Barcelona, where many firm headquarters are located. Although our database might produce these mis-allocations, they do not invalidate the results from empirical tests on export specialization as long as the probability of a mis-allocation is the same across industries. We do not find any reason to expect some industries exports to have a higher probability of being mis-classified than others. To test the robustness of our results, we also perform the empirical tests excluding the urban areas of Madrid and Barcelona, which are the most likely to experience the headquarter effect. Results are not altered.

We remove from the database marginal export operations (those with a value below 1,500 euros) and firms operating in the energy sector. Due to their special circumstances, we exclude export operations to Andorra and Gibraltar. Finally, we remove from the sample the urban areas that report very few exports operations (less than 20). This leaves a final database with 64 urban areas.[5]

We also use alternative export data sources to test the robustness of our results. First, the SABI database, produced by Bureau van Dijk, provides economic and financial information for around 2 million Spanish firms. SABI identifies the municipality in which the firm is located, whether the firm exports or not, and in latest releases, the value of exports as % of total sales.[6] We calculate urban areas exports multiplying the share of exports by total sales and aggregating municipalities into urban areas. SABI data does not provide exports data by destination, so we can only carry out the empirical tests using exports by urban area and industry. Second, we use province level (NUTS-3) international trade data from the Inland Revenue Agency. This database provides the value of exports per province, 8-digit European Common Nomenclature and destination. The main limitation of this data is that Spanish provinces might encompass more than one metropolitan area, or include municipalities that do not belong to any urban area. To determine whether province level international trade data is representative of the an urban area trade pattern, we calculate the share of the urban area exports in total province exports using SABI data. If the urban area represents more than two-thirds of province exports, we consider that province exports are representative of the urban area trade pattern.

[5]The urban areas excluded from the sample are Algeciras, Arrecife, Benidorm, Cáceres, Ceuta, Fuengirola, La Línea de la Concepción, Marbella, Melilla, Santa Lucía de Tirajana, Torrevieja, and Zamora.

[6]The SABI database might also present the mis-allocation problem explained before.

*Industries' skill intensity, complexity, and skill types*

We proxy industries' skill intensities by the share of employees in skilled occupations over total employees. We consider skilled occupations those included between the Standard Occupational Classification (SOC) categories 11 and 29: management and other occupations that involve an intensive use of scientific and technical knowledge. This data is obtained from the Occupational Employment Statistics (OES) survey of the U.S. Bureau of Labor Statistics (available at `http://www.bls.gov/oes`). As an alternative skill intensity measure, following Romalis (2004) and Chor (2010), we calculate the share of non-production workers over total workforce. This data is obtained from the US Manufacturing Census.[7]

To calculate industries intensity in cognitive, people and connected skills we use the O*NET database (available at `http://www.onetcenter.org/database.html`). O*NET identifies 21 abilities that are related to the cognitive area and weights their importance for each occupation (in a 1 to 5 scale).[8] Following Bacolod et al. (2009), we use a principal component analysis to collapse the 21 abilities into one cognitive skill indicator. We calculate an industry-level cognitive skill intensity as the sum of all occupations cognitive skill intensity, weighted by the share of each occupation in total employment. We obtain this latter data from OES. We use a similar procedure to calculate the industry level intensity of social skills. In this case the O*NET identifies six different social skills.[9] Finally, we use the Kok and Weel (2014) connectivity index to calculate the connectivity intensity in each industry. These authors calculate the spatial correlation of the work activities identified in O*NET (see Table 2 in Kok and Weel (2014)). From these data we calculate the connectivity of each occupation as the average of each work activity, weighted by the importance of the activity in each occupation. Then, we calculate an industry level skills connectivity index as the average of occupations connectivity, weighted by the share of each occupation in employment.

We use two different variables to measure the diversity of skills used in an industry, or industry complexity. The first is obtained from Hausmann and Hidalgo (2014), which define industry level complexity as the number of capabilities that are needed to produce a good. The number of capabilities is proxy through a iterative process between countries diversification level (number of industries that countries export with a revealed comparative advantage) and industries ubiquity (number of countries that have

[7]We also use other proxies for skill intensity, such as the share of non-production workers wages in total payroll and average wage, with no change in results.

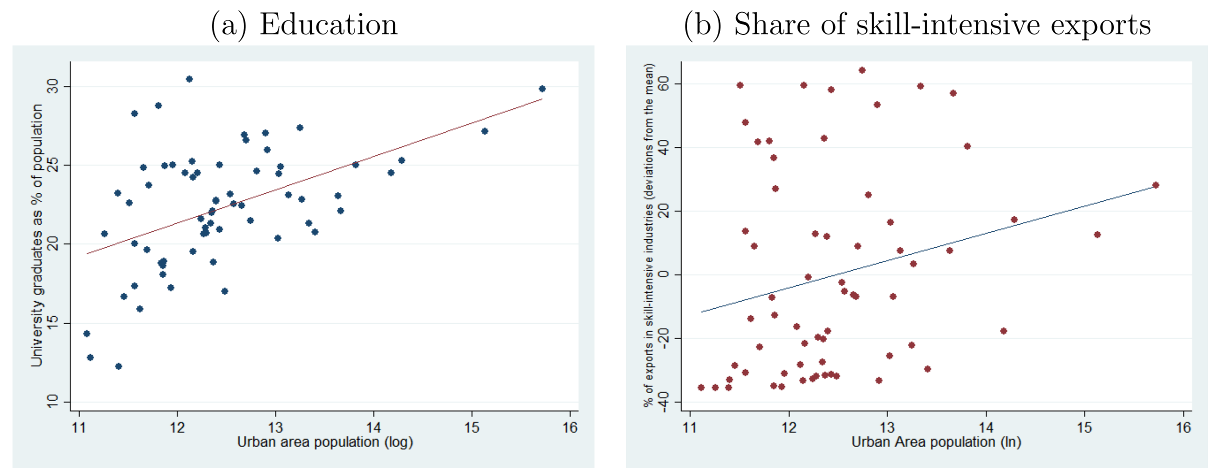
[8]The 21 abilities are category flexibility, deductive reasoning, flexibility of closure, fluency of ideas, inductive reasoning, information ordering, mathematical reasoning, memorization, number facility, oral comprehension, oral expression, originality, perceptual speed, problem sensitivity, selective attention, spatial orientation, speed of closure, time sharing, visualization, written comprehension and expression.

[9]These are coordination, instructing, negotiation, persuasion, service orientation and social perceptiveness.

a revealed comparative advantage in the industry). This data, calculated at the Harmonized System 4-digit level, is available at `http://atlas.cid.harvard.edu/rankings/`. The second variable comes from Minondo and Requena-Silvente (2013), which define complexity as the number of skill-intensive tasks that are combined to produce a good. Skill-intensive tasks are defined in the same way as for the skill intensity variable. The data to calculate this latter variable is obtained from OES.[10]

To finish this section, we present two figures to motivate the empirical tests. Figure 2, Panel a, shows the relationship between urban areas population and the share of university graduates in total population. The figure shows a strong positive relationship between both variables. This result is in line with previous studies, such as Glaeser and Resseger (2010) and Davis and Dingel (2014) for the US. De la Roca and Puga (2012) also show a very strong positive relationship between city size and mean annual earning in Spanish urban areas.

**Figure 2:** Correlation between populations, education and skill-intensive exports, 2012




Source: Own elaboration from INE Census and Directory of Exporting and Importing Firms.

As argued in the theoretical section the concentration of skill workers in large cities should render them a comparative advantage in skill-intensive industries. Panel (b) shows the relationship between the share of skill-intensive exports and urban area population. Despite some dispersion, we observe an overall positive relationship between the urban area population and the share of skill-intensive exports. Next section presents different

[10]As cognitive skills, social skills, connectivity and Minondo-Requena complexity variables can only be calculated with US sources, for the sake of consistency, we also use a US source to calculate the skill intensity variable. To test the robustness of our results, we also calculate a skill-intensity measure using Spanish data. In particular, we use INE's Industrial Survey to calculate wages per hour and wages per employee by industry. Empirical results are not altered.

tests to confirm this relationship.

# 4 Results

In this section we present the results of the empirical tests. The section is divided in three parts. First, we perform the empirical tests using exports data at the urban area and industry level. Second, we extend the analysis to export data by urban area, industry and destination. Finally, we carry out some additional analyses to test the robustness of our results.

## 4.1 Empirical tests with exports by urban area and industry

The first empirical test on the trade pattern of cities is the elasticity-test. If exports are log-supermodular in population and skill intensity, then the population elasticity of exports should increase in industries' skill intensity. To perform this empirical test we estimate the following regression equation:

$$\ln x_i^k = \alpha^k + \beta_k' \alpha^k \ln pop_i + \epsilon_i^k \tag{2}$$

where $\alpha^k$ is a dummy variable if exports correspond to industry $k$ and zero otherwise, and $\beta_k'$ is a vector of industry-specific population elasticities.

**Figure 3:** Industries' population elasticities and skill intensities

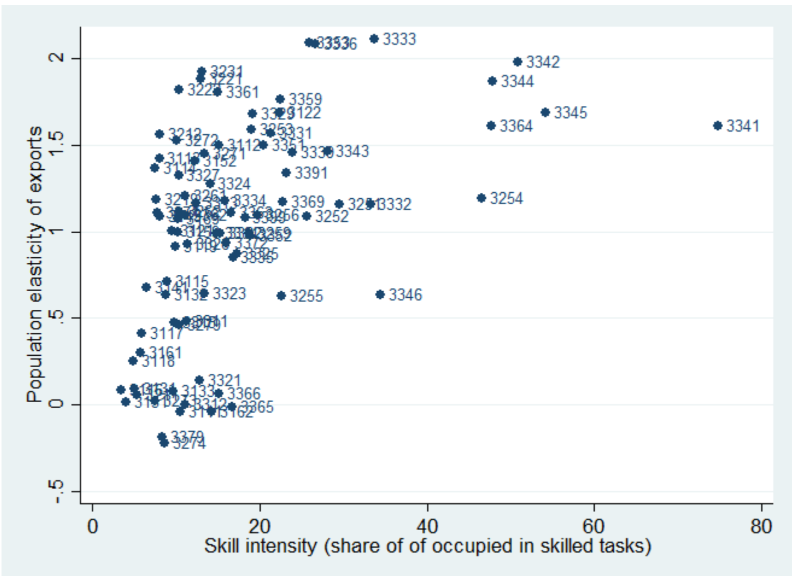



Figure 3 shows the relationship between the population elasticity of exports and industries' skill intensity, measured as the share of employees in skilled occupations over total employees. Among the 84 NAICS 4-digit industries included in the graph, computer and peripheral equipment manufacturing (code 3341) is the most skill-intensive industry. The figure shows a strong positive relationship between industries skill intensity and population elasticities. In fact, the empirical test concludes that more skill-intensive industries have higher elasticities in 97% of comparisons.[11] The rate of success is higher than the one reported by Davis and Dingel (2014) using employment data.

Figure 4 presents the correlation between industries' population elasticities and other industry-level skills and complexity measures. In Panel (a) we introduce an alternative measure for industries skill intensity: the share of non-production employees over total employees; Panel (b) and (c) show the relationship for industries intensity in social and cognitive skills respectively; Panel (d) shows industries' intensity in tasks that require physical proximity; finally, Panel (e) and (f) show the relationship between elasticities and industries' complexity, measured by the Hausmann-Hidalgo and the Minondo-Requena indicators. In all cases, there is a positive relationship between industries skill/complexity intensity and population elasticities. The empirical test confirms the hypothesis for all variables for, at least, 94% of cases.[12]

The second empirical test proposed by Davis and Dingel (2014) is the pairwise comparisons test. If exports are log-supermodular in cities population and industries skill intensity, for any pair of industries and cities the share of exports of the more skill-intensive industry over the less skill-intensive industry should be equal or larger in the more populated city than in the less populated city. Following Davis and Dingel (2014), to perform this test we aggregate urban areas into bins according to population size. For example, in the first pairwise comparison urban areas are aggregated in two bins: the large-population bin and the small-population bin. Afterwards, we repeat the test for an increasing number of bins, until we reach the highest disaggregation, where the number of bins equals the number of urban areas in the sample. Obviously, as the number of bins increases the test performs a larger number of comparisons.

Figure 5 shows the results of the pairwise comparisons. The horizontal axis measures the number of bins that are considered in the analysis and the number of comparisons made; the vertical axis measures the percentage of times the theoretical prediction is validated. The triangles show the success rate when all comparisons have the same weight (unweighted), and the diamonds show the success rate when comparisons are weighted by differences in population and industries skill intensity.[13]

[11] With 84 industries, we perform (84*83)/2=3,486 comparisons, in which 3,393 of them the hypothesis is confirmed. We use the 5% threshold to test the validity of the null hypotheses.

[12] The hypothesys is confirmed in 94% of cases for the share of non-production workers, 96% for social skills, 97% for cognitive skills and connectivity and 95% for the two complexity indicators.

[13] As explained by Davis and Dingel (2014), it is sensible to weight the success rate by the differences in population between cities and skill intensity between industries. For example, let's consider exports

**Figure 4:** Correlation of industries' population elasticities and other skills and complexity measures, 1975–2009

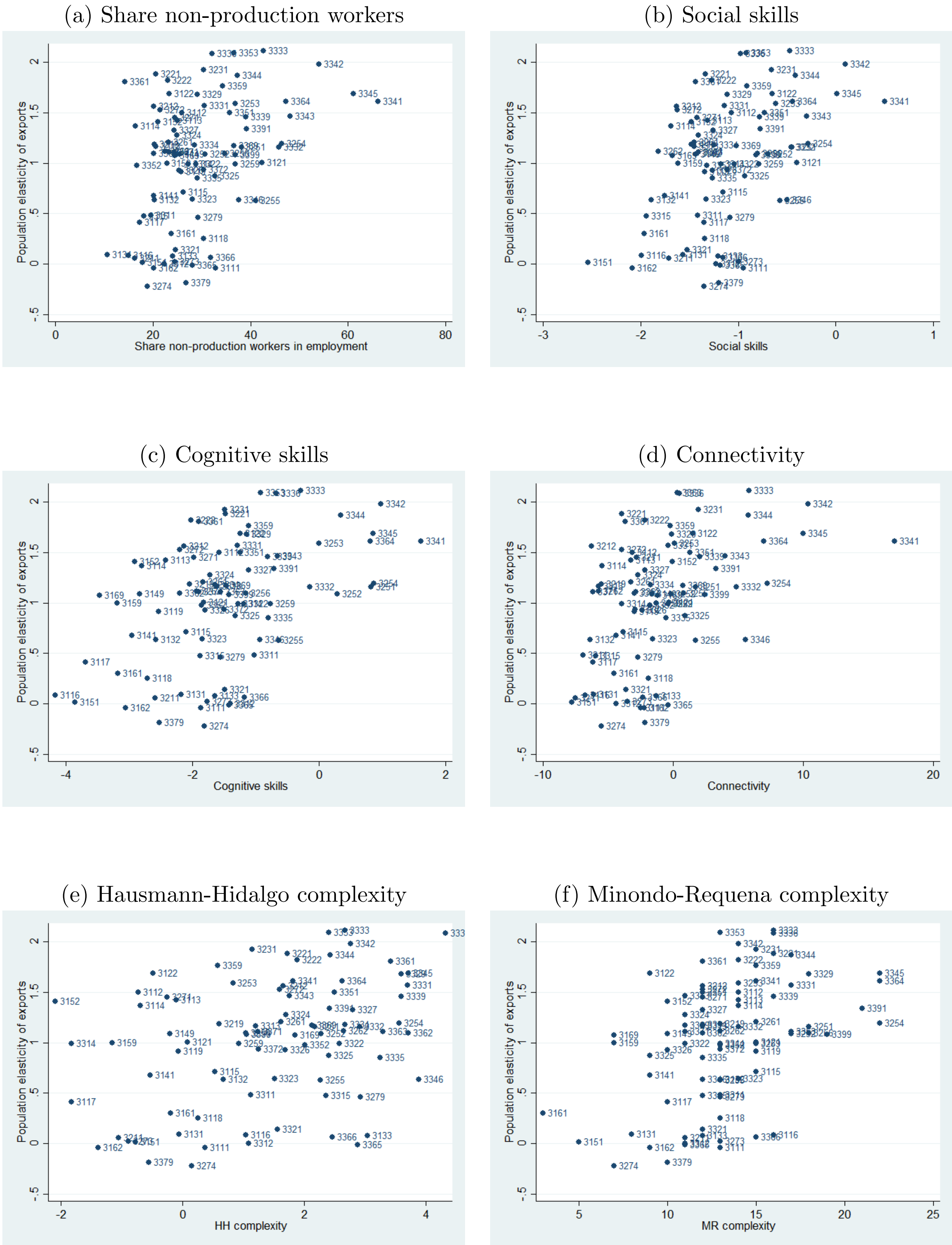


**Figure 5:** Pairwise comparisons (relative skill intensity and export shares)


The sample includes 84 (4-digit NAICS) industries and 64 urban areas. We only carry out pairwise comparisons when the share of exports of an urban area in a given industry is different from zero.

When the urban areas are divided in two groups, 3,403 comparisons are made, and the theoretical prediction is validated in 63% of the unweighted cases and 73% of the weighted cases.[14] As the number of bins increases there is a reduction in the success rate. When all urban areas are compared the percentage of success is 53% for the unweighted cases and 59% for the weighted cases. These percentages validate the theoretical prediction that larger cities export more skill-intensive goods. Our results are very similar to those reported by Davis and Dingel (2014) using employment data.

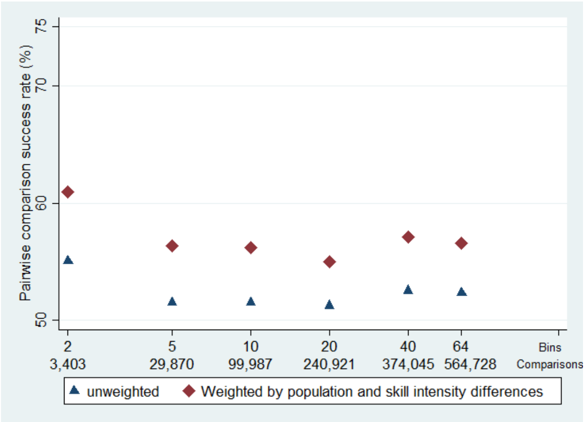
Figure 6 presents the results of the pairwise comparison tests for other skill and complexity indicators. In all cases the highest success rate is achieved when urban areas are divided in two bins. In Panel (a) we present the success rate for the alternative

of computer and peripheral equipment (the most skill-intensive industry) relative to fiber, yarn, and thread mills (the least skill-intensive industry). The finding that this ratio is higher in Girona (the least populated urban area) than in Madrid (the most populated urban area) would be more damaging to the theoretical model than the finding that the ratio is higher in Girona than in Santiago de Compostela (the second least populated urban area). Similarly, the damage to the model would be lower if the exports ratio referred to the second-least intensive industry relative to the least skill-intensive industry, rather than the ratio of the most skill-intensive industry relative to the least skill-intensive industry. To address these effects, each comparison is weighted by the product of the (log) difference in population and the difference in skill intensity.

[14] The maximum number of pairwise comparisons for two bins would be = (84 industries * 83 industries) /2 = 3,486. If we omit the 83 comparisons where at least a trade share was zero, we end-up with 3,403 comparisons.

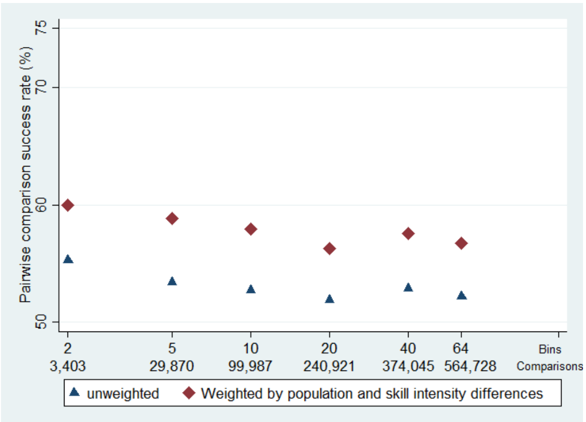
**Figure 6:** Pairwise comparisons. Other skill and complexity indicators

(a) Share non-production workers

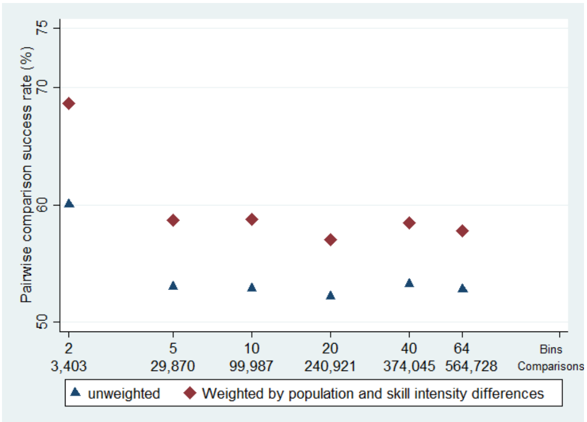



(b) Social skills

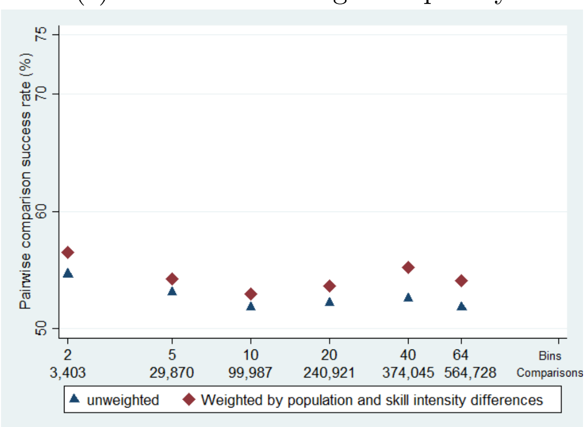

(c) Cognitive skills



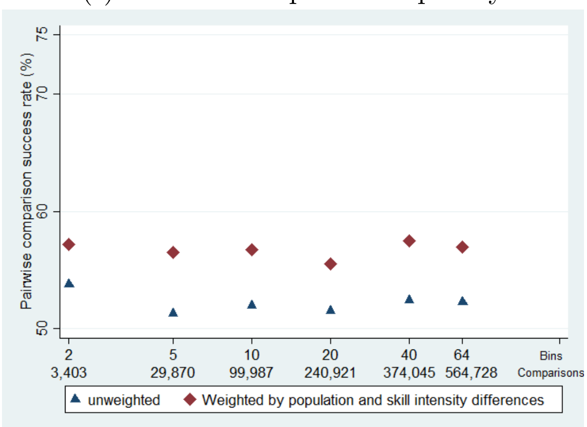

(d) Connectivity




(e) Hausmann-Hidalgo complexity




(f) Minondo-Requena complexity



measure of skill-intensity: the share of non-production workers in the labor force. The maximum success is 60%, when urban areas are divided in two bins and comparisons are weighted by skill and population differences. The maximum success rates are 62% and 60% for social and cognitive skills respectively (Panels b and c). The maximum success rate is higher for connectivity-intensive industries: 68% (Panel d). Regarding complexity measures, the maximum success rate is around 57% for both indicators (Panels e and f). These results confirm that most populated cities export relatively more in industries that make intensive use of social and cognitive skills, and tasks that are performed at a close distance. Highly populated cities also export relatively more in industries that require the combination of a large number of different skills.

## 4.2 Empirical tests with exports by urban area, industry and destination

In this subsection, we take advantage of the destination dimension of our export data to estimate equation (1). As mentioned above, in this equation cities trade specialization is captured by the interaction term between the industry skill intensity and the urban area population.

Table 1 presents the results of the estimation. In column (1) we interact the urban area population with the industry skill intensity, measured as the share of employees in skilled tasks over all employees. The coefficient is positive and strongly statistically significant. This result confirms that urban areas with a large population export skill-intensive products. As explained in the theoretical section, the positive relationship between city size and skill-intensive exports arises because the concentration of skilled workers in large urban areas, and the complementarity between skilled workers and skill-intensive goods. To confirm this hypothesis, in column (2), we interact industries skill-intensity with urban areas share of university graduates. As expected, the interaction coefficient is positive and statistically significant. It would be interesting to put in the same regression the interaction coefficients of column (1) and (2); however, due to the very high correlation between both interaction coefficients, estimations would not be valid. There is also a very high correlation between the skill intensity variable and the cognitive, social and connectivity variables. Hence, we also run separate regressions for these variables. In columns (3) and (4) we analyze whether cities specialize in industries that make intensive use of social and cognitive skills. We find that in both cases the interaction coefficient is positive and statistically significant. We also find that cities export in industries that demand tasks that should be performed in geographical proximity (column 5). Finally, the interaction term between the Hidalgo-Hausmann complexity variable and urban area population, and the interaction term between the Minondo-Requena complexity variable

**Table 1:** Cities population and export specialization in skill-intensive and complex industries

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Share skilled occup * ln (pop) | 0.0121*** | | | | | | | 0.00272** |
| | (0.001) | | | | | | | (0.001) |
| Share skilled occup * share educ | | 0.00406*** | | | | | | |
| | | (0.001) | | | | | | |
| Social skills * ln(pop) | | | 0.381*** | | | | | |
| | | | (0.031) | | | | | |
| Cognitive skills * ln(pop) | | | | 0.161*** | | | | |
| | | | | (0.015) | | | | |
| Connectivity * ln(pop) | | | | | 0.0353*** | | | |
| | | | | | (0.004) | | | |
| HH complexity * ln(pop) | | | | | | 0.0916*** | | 0.0157 |
| | | | | | | (0.010) | | (0.011) |
| MR complexity * ln(pop) | | | | | | | 0.0704*** | 0.0625*** |
| | | | | | | | (0.004) | (0.004) |
| Observations | 30657 | 30657 | 30657 | 30657 | 30657 | 30657 | 30657 | 30657 |
| $R^2$ | 0.515 | 0.513 | 0.518 | 0.517 | 0.514 | 0.513 | 0.523 | 0.523 |

Note: all regressions include city+country of destination and country of destination+industry fixed effects. Clustered standard errors at the origin city and destination country combination in parentheses. ***, ** statistically significant at 1% and 5% respectively.

and urban area population are also positive and statistically significant. These results point out that larger cities specialize in the exports of goods that demand a large combination of different skills. Finally, we introduce in the same regression the skill intensity variable and the two complexity variables. We find that the skill-intensity variable and the Minondo-Requena complexity variables remain positive and statistically significant. This result suggest that large cities have a comparative advantage in industries intensive in skills and that require the combination of a large number of skilled tasks.

## 4.3 Robustness

In this subsection we analyze whether our empirical results are robust to alternative samples, exports data sources and urban area classifications. First, we analyze whether our results are robust to excluding Madrid and Barcelona, the two largest urban areas, from the sample. These urban areas are the most likely to experience a headquarter effect, and we want to test that they are not biasing our results. Figure A1, Panel a, shows the relationship between the population elasticity of exports and the industry skill intensity. There is a positive relationship between both variables and the elasticity-test is validated in 92% of cases. This success rate is only 5 percentage points lower than in the baseline sample. Figure A2 shows the results of the pairwise test. The (weighted) success rate is 64% when urban areas are divided in two bins and drops to 54% when urban areas are disaggregated at the highest level (62). These success rates are lower than in the baseline sample (73% and 59%), but still above 50%. Finally, Column 1 of Table A2 shows the estimation of the interaction term between industries skill-intensity and urban areas population. The coefficient remains positive and statistically significant. However, the size of the coefficient is reduced by half when compared with the baseline estimate (0.005 vs. 0.0121).

Second, as mentioned in the Data section, we carry out the empirical tests using alternative sources for urban areas' export data: SABI and Inland Revenue data for provinces. SABI only allows to calculate export data by urban area and industry, so we only carry out the elasticity test and the pairwise comparison test. Figure A1, Panel b, shows the relationship between the population elasticity of exports and industries skill intensity. Although there are some outliers, the elasticity test is validated in 83% of cases. The pairwise comparison yields a maximum 73% success rate when urban areas are divided in two are comparisons are weighted by population and skill intensity differences (Figure A2, Panel b).[15] We also use province level data from the Inland Revenue Agency to proxy urban area exports. To determine whether province level international trade data is representative of the an urban area trade pattern, we calculate the share of

[15]Note that the number of comparisons is much lower than in the baseline case as there is a larger number of zeros per NAICS 4-digit industries and urban areas in the SABI database.

the urban area exports in total province exports using SABI data. If the urban area represents more than two-thirds of province exports, we consider that province exports are representative of the urban area trade pattern.[16] There are 18 urban areas that meet the 2/3 threshold.[17] The elasticity test is validated in 94% of cases (Figure A1, Panel c) and the pairwise comparison test yields a maximum 70% success rate (Figure A2, Panel c). The econometric analysis also confirms a positive relationship between urban area population and exports and skill intensive goods (Table A2, Column 2).[18]. The interaction coefficient is larger than in the baseline scenario.

Finally, we use an alternative classification of urban areas to test the robustness of our results. The alternative classification is the one produced by the Spanish Ministry of Public works, as described in the Data section. This classification has 86 urban areas, 10 more than the OECD classification. Comparing both classifications, they have 573 municipalities in common, there are 1,844 municipalities that appear only in the OECD database, and 186 municipalities areas that appear only in the Ministry of Public Works classification. The results of the empirical tests are very similar to those of the OECD urban area sample.

# 5 Conclusions

As Behrens and Robert-Nicoud (2015) highlight, cities differ in many ways. In this paper we explore whether large cities differ from small cities in trade patterns. Using exports data for the Spanish urban areas for the year 2012, we show that large cities export more skill-intensive goods than smaller cities. Our empirical tests also conclude that large cities export goods that are more intensive in social and cognitive skills, that demand tasks that should be performed at a geographical proximity, and demand a large range of skills. Our empirical results are robust to removing the largest urban areas, the use of alternative export data sources, and the use of alternative urban area classifications.

# References

Bacolod, M., Blum, B. S., and Strange, W. C. (2009). Skills in the city. *Journal of Urban Economics*, 65(2):136–153.

Behrens, K. and Robert-Nicoud, F. (2015). Agglomeration theory with heterogeneous

[16] Results are robust to the use of alternative thresholds (50% and 75%) and criteria (population and GDP of the urban area as a share of the province).

[17] Albacete, Barcelona, Bilbao, Cádiz, Córdoba, Illes Balears, La Rioja, Las Palmas, León, Madrid, Ourense, Pamplona, Salamanca, Santa Cruz de Tenerife, Sevilla, Valencia, Valladolid and Vigo.

[18] Note that the number of observations is larger in the regression. Despite the lower number of urban areas, province-level data has a lower number of zeros per industry and destination

agents. in G. Duranton, J.V. Henderson, and W. Strange (eds.) *Handbook of Regional and Urban Economics, Volume 5.* Amsterdam: Elsevier.

Chor, D. (2010). Unpacking sources of comparative advantage: A quantitative approach. *Journal of International Economics*, 82(2):152–167.

Costinot, A. (2009a). An elementary theory of comparative advantage. *Econometrica*, 77(4):1165–1192.

Costinot, A. (2009b). On the origins of comparative advantage. *Journal of International Economics*, 77(2):255–264.

Davis, D. R. and Dingel, J. I. (2012). A spatial knowledge economy. Working Paper 18188, National Bureau of Economic Research.

Davis, D. R. and Dingel, J. I. (2014). The Comparative Advantage of Cities. Working Paper 20602, National Bureau of Economic Research.

De la Roca, J. and Puga, D. (2012). Learning by working in big cities. *CEPR Discussion Paper No. DP9243*.

Duranton, G. and Puga, D. (2001). Nursery cities: Urban diversity, process innovation, and the life cycle of products. *American Economic Review*, 91(5):1454–1477.

Gaspar, J. and Glaeser, E. L. (1998). Information technology and the future of cities. *Journal of urban economics*, 43(1):136–156.

Glaeser, E. L. (1999). Learning in cities. *Journal of urban Economics*, 46(2):254–277.

Glaeser, E. L. (2011). *Triumph of the city: How our greatest invention makes US richer, smarter, greener, healthier and happier.* Pan Macmillan.

Glaeser, E. L. and Resseger, M. G. (2010). The complementarity between cities and skills*. *Journal of Regional Science*, 50(1):221–244.

Hausmann, R. and Hidalgo, C. A. (2014). *The atlas of economic complexity: Mapping paths to prosperity.* MIT Press.

Hausmann, R., Hwang, J., and Rodrik, D. (2007). What you export matters. *Journal of economic growth*, 12(1):1–25.

Jacobs, J. (1969). *The economy of cities.* Random House, New York.

Kok, S. and Weel, B. t. (2014). Cities, tasks, and skills. *Journal of Regional Science*, 54(5):856–892.

Minondo, A. and Requena-Silvente, F. (2013). Does complexity explain the structure of trade? *Canadian Journal of Economics/Revue canadienne d'économique*, 46(3):928–955.

OECD (2012). *Redefining Urban: A New Way to Measure Metropolitan Areas*. OECD Publishing.

Romalis, J. (2004). Factor Proportions and the Structure of Commodity Trade. *American Economic Review*, 94(1):67–97.

**Table A1:** Urban areas in Spain. OECD classification

| Name | Type | Population (2012) |
|---|---|---|
| Madrid | Large | 6,594,797 |
| Barcelona | Large | 3,703,871 |
| Valencia | Metropolitan | 1,534,540 |
| Sevilla | Metropolitan | 1,446,746 |
| Bilbao | Metropolitan | 990,908 |
| Zaragoza | Metropolitan | 827,178 |
| Málaga | Metropolitan | 790,450 |
| Las Palmas | Metropolitan | 657,053 |
| Mallorca | Medium-size | 620,907 |
| Murcia | Medium-size | 577,024 |
| Granada | Medium-size | 519,423 |
| Tenerife | Medium-size | 506,612 |
| Vigo | Medium-size | 453,342 |
| Valladolid | Medium-size | 448,367 |
| Alicante | Medium-size | 445,838 |
| A Coruña | Medium-size | 409,392 |
| Córdoba | Medium-size | 363,813 |
| Pamplona | Medium-size | 350,782 |
| Cádiz | Medium-size | 343,197 |
| Oviedo | Medium-size | 322,836 |
| Santander | Medium-size | 314,446 |
| Gijón | Medium-size | 288,700 |
| Vitoria | Medium-size | 272,658 |
| Elche | Medium-size | 264,552 |
| San Sebastián | Medium-size | 260,131 |
| Marbella | Small | 249,148 |
| León | Medium-size | 244,984 |
| Terrassa | Small | 240,932 |
| Almería | Medium-size | 237,896 |
| Cartagena | Small | 235,664 |
| Castellón | Small | 233,426 |
| Sabadell | Medium-size | 231,301 |
| Salamanca | Medium-size | 223,171 |
| Jerez | Small | 218,991 |
| Lleida | Small | 212,501 |
| Logroño | Small | 200,922 |
| Albacete | Small | 198,760 |
| Burgos | Medium-size | 198,300 |

| Name | Type | Population(2012) |
|---|---|---|
| Huelva | Small | 193,203 |
| Tarragona | Small | 189,724 |
| Badajoz | Small | 188,030 |
| Ourense | Small | 187,659 |
| Santiago de Compostela | Small | 180,558 |
| Girona | Small | 174,754 |
| Benidorm | Small | 159,785 |
| Fuengirola | Small | 158,077 |
| Pontevedra | Small | 155,955 |
| Ferrol | Small | 151,888 |
| Jaén | Small | 143,348 |
| Arrecife | Small | 142,132 |
| Mataró | Small | 141,825 |
| Algeciras | Small | 140,058 |
| Avilés | Small | 140,033 |
| Reus | Small | 137,600 |
| Manresa | Small | 133,033 |
| Toledo | Small | 126,737 |
| Cáceres | Small | 124,518 |
| Lugo | Small | 119,498 |
| Gandia | Small | 114,544 |
| Talavera | Small | 111,028 |
| Guadalajara | Small | 109,684 |
| Torrevieja | Small | 108,923 |
| Granollers | Small | 105,460 |
| Ciudad Real | Small | 102,641 |
| Vilanova | Small | 100,015 |
| Palencia | Small | 97,935 |
| Elda | Small | 89,233 |
| Pto. Sta. María | Small | 89,068 |
| Ponferrada | Small | 87,973 |
| Zamora | Small | 87,898 |
| Ceuta | Small | 84,018 |
| Melilla | Small | 80,802 |
| Irun | Small | 77,620 |
| Sanlúcar de Barrameda | Small | 67,308 |
| Sta. Lucía de Tirajana | Small | 67,291 |
| La Línea de la Concepción | Small | 64,704 |

* Source: http://www.oecd.org/gov/regional-policy/all.pdf and INE.

**Figure A1:** Elasticity test. Robustness


(a) Excluding Madrid+Barcelona

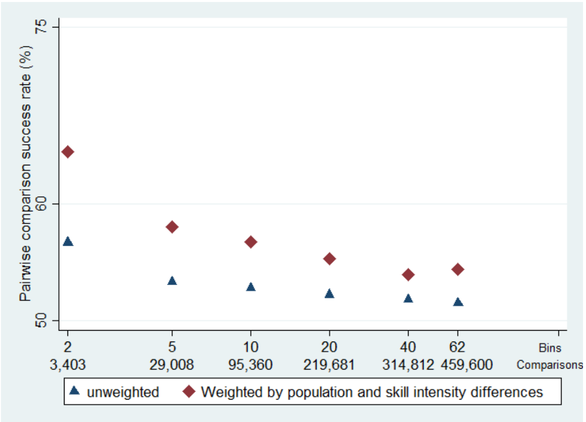
(b) SABI data

(c) Province data

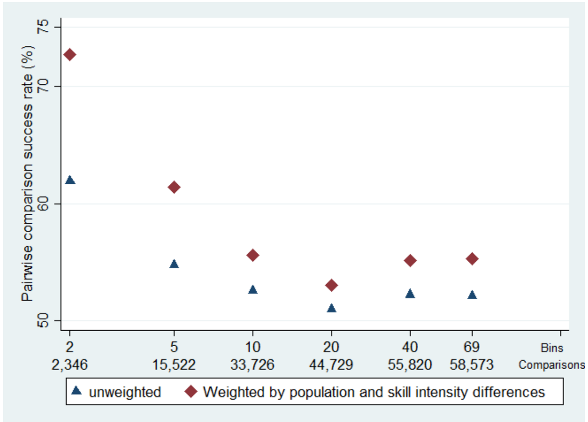
(d) Alternative urban areas

**Figure A2:** Pairwise comparisons. Robustness

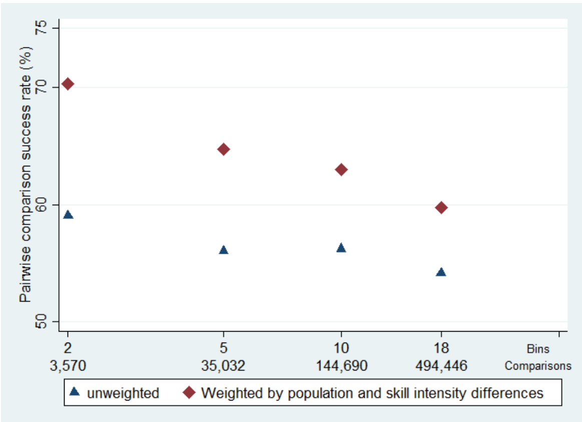
(a) Excluding Madrid+Barcelona




(b) SABI data




(c) Province data




(d) Alternative urban areas

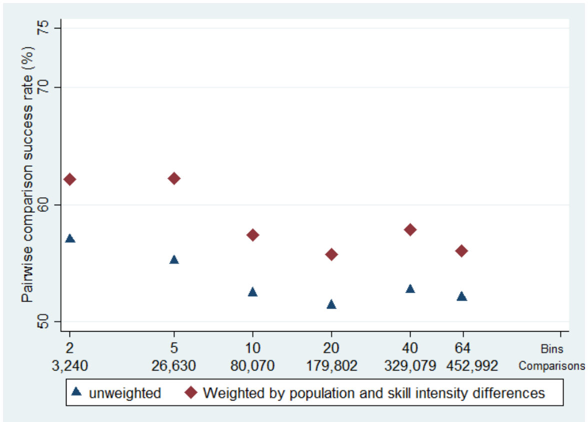



**Table A2:** Robustness analysis. Cities population and export specialization in skill-intensive and complex industries.

| | (1) No MAD-BCN | (2) Province data | (3) Alternative UA |
|---|---|---|---|
| Share skilled occup * ln (pop) | 0.00541** | 0.0173*** | 0.0118*** |
| | (0.002) | (0.001) | (0.001) |
| Observations | 22227 | 60117 | 29115 |
| $R^2$ | 0.512 | 0.574 | 0.531 |

Note: all regressions include city+country of destination and country of destination+industry fixed effects. Clustered standard errors at the origin city and destination country combination in parentheses. ***, ** statistically significant at 1% and 5% respectively.