WORKING PAPERS

The Fickle Fringe and the Stable Core: Exporters' Product Mix Across Markets

Lionel Fontagné
Angelo Secchi
Chiara Tomasi

CESifo
Center for Economic Studies & Ifo Institute

The Fickle Fringe and the Stable Core: Exporters' Product Mix Across Markets

Abstract

Multi-product exporters choose their product mix focusing on their best-performing products. Although their product mix varies across countries (the fickle fringe), the interdependence in demand or production technology making vectors of products systematically co-exported leads to commonalities in this mix across destinations (the stable core). In order to uncover the determinants of the fickle and stable parts of firm export product mix, we use a cross section of firm-product-destination level French and Italian data, taking explicitly into account the choice of not exporting a product to a destination. Using dissimilarity measures instead of rank correlations, we observe a great deal of variability among the product-mixes a firm exports to different destinations. We show that market size, but also the market positioning of a firm and market structure explain part of this observed variability. At the same time, together with this fickle part, we highlight the existence of a stable component among the diverse product-mixes exported. The probability of exporting this core set of products increases with the size of the destination market and with the ability to match demand, but is inversely related to market concentration.

Lionel Fontagné
PSE – University of Paris 1
Paris / France
lionel.fontagne@univ-paris1.fr

Angelo Secchi
PSE – University of Paris 1
Paris / France
angelo.secchi@univ-paris1.fr

Chiara Tomasi
University of Trento
Trento / Italy
chiara.tomasi@unitn.it

1 Introduction

The typical distribution of exports is skewed: across 32 developing countries, the top five exporters make up 30% of total exports (Freund and Pierola, 2015). Such large multi-product and multi-destination exporters also appear in emerging or developed countries: in Brazil 25% of exporters ship ten or more HS6 products and account for 75% of total exports (Arkolakis and Muendler, 2010), while in the US firms shipping more than five HS10 products represent 30% of total exporters and account for 97% of all exports (Bernard et al., 2009). Similar figures are observed in Europe: in Italy and France 42% and 40% of exporters ship more than five HS6 products and account for 96% and 95% of total export flows, respectively. Accordingly, an overwhelming share of international trade is conducted by large firms that export a broad variety of products to different destinations.

The selection of *destinations* by these exporters is predicted by models with heterogeneous firms and asymmetric countries. Firms in a given exporting country face a list of destinations by decreasing order of accessibility: only the most efficient exporters will export to the most difficult destinations (Chaney, 2008; Eaton et al., 2011). The selection of the *products* exported can first be predicted from firm capabilities, assuming a product ladder where productivity falls for each additional variety produced. For example Eckel and Neary (2010) consider a flexible manufacturing approach, where firms face declining efficiency in supplying less-successful products away from their core competency. In the same model, on the demand side, new varieties reduce the demand for the firm's existing varieties.[1] Alternatively, Mayer et al. (2014) assume firm-product specific marginal costs that increase in scope together with non-CES preferences, which generates the prediction that firm sales will be more skewed towards their core competency in more competitive markets.

Both of these approaches shed light on one dimension of the complex problem faced by multi-

---

[1]Another class of multi-product firm models in international trade assumes that products are symmetric on both the demand and supply sides. As a result, the same amount of all products is sold. See, e.g. Nocke and Yeaple (2014); Feenstra and Ma (2008); Dhingra (2013).

product exporters: cannibalization across own exported products in Eckel and Neary (2010) and sales that are skewed towards the firm's better-performing products in more difficult markets in Mayer et al. (2014).[2] The focus on cannibalization predicts a change in the list of exported products at the firm-product extensive margin when fully opening to trade with the rest of the world, while focusing on differences in mark-ups across asymmetric destinations mainly leads to adjustments at the firm-product intensive margin for a given product-mix.[3] A common feature of all these models is that firms enter export markets with their most efficient product first, and then expand their scope moving up the marginal-cost ladder. As a consequence, a firm's successful products in one destination should also be their leading products in all of the firm's other markets, with the same hierarchy of products being observed across the different destinations.

Simple rank correlations show however that exporter product mix is neither strictly ordered across destinations, nor totally idiosyncratic: firm product sales are only *imperfectly* correlated across markets.[4] The hierarchy of products across destinations is not fixed as certain firm-product attributes are potentially idiosyncratic across export destinations. The stochastic model in Bernard et al. (2011), where product attributes vary by both products and countries within a firm, yields this kind of heterogeneity in firm product mix across destinations. This rank-correlation evidence suggests that although there are some commonalities in the list of products exported, with there being a *stable core*, there are also widespread and systematic differences across destinations, producing a *fickle fringe* in the firm export product mix.

---

[2]In Arkolakis and Muendler (2010) firm sales within a destination are concentrated on a few core products, wide-scope exporters sell more of their top selling products than do narrow-scope exporters, and sales of the lowest-selling products fall with scope. This prediction results from the combination of product-specific marginal costs that fall in the firm's scope, with product-destination specific local entry costs increasing in the number of products exported.

[3]Adjustment at the firm-product extensive margin also appears in Mayer et al. (2014), but is sensitive to the model's assumptions regarding fixed costs.

[4]Mayer et al. (2014) find a rank correlation coefficient between French firms' local and global product rank of 0.68. Arkolakis and Muendler (2010) calculate the correlation in Brazilian firm-product sales' ranking by destination, using either the US or Argentina as the reference country, and find an average coefficient of 0.85. Di Comite et al. (2014) find an average pairwise bilateral correlation of 0.50 over a set of countries in the country-pair of Belgian firm-product exports. Similar conclusions are reached in empirical analyses looking at firm sales variation in different countries (Eaton et al., 2011; Kee and Krishna, 2008).

How common are the coexisting fickle and stable components in exporting firms' product mix? And what are the main factors driving this variation in product mix across destination markets?

We use a cross section of French and Italian data at the firm-product-destination level, and provide a novel analysis of the export patterns of multi-product firms across destination markets. We measure the extent of this variability and identify the determinants and properties of both the fickle fringe and the stable core of firm export mix. While the literature considers ranked lists of independent products, we innovate by shifting the focus towards mixes of products that are jointly exported. By doing so, we account for the interdependence in demand or production technology that makes product pairs (or more generally *vectors*) systematically co-exported. We also explicitly control for the firm's choice of *not* exporting some of its products.

We find a great deal of variation in firm export product-mixes to different destinations, both in terms of the combination of products and their relative importance as measured by export value. A firm-destination measure of market size plays a major role in explaining this variability, as do firm-destination market structure and the ability to match demand. At the same time, together with this fickle fringe, there is a stable product-mix component. We show that this core combination of products is both exported frequently and accounts for considerable export value. We also show that some items in this stable core would not have been identified using the criterion of sales value. In line with economic intuition, we show that the probability of exporting this core set of products rises with the size of the destination market and the ability to match demand, but falls with market concentration.

We interpret our results as revealing widespread production and demand complementarities, making some firm product combinations more likely to be exported than others:[5] firms are multi-product not only to diversify their risks, but also for technological and demand reasons. For the

---

[5]This result is in line with the finding in Bernard et al. (2010) on domestic production using US Manufacturing Census data.

former, economies of scope generated by excess capacity in market and internal productive factors, including organizational knowledge, are traditionally considered as the main motivation for firm product diversification.[6] For the latter, complementarities and strategic product bundling/tying to exploit price discrimination or leverage market power across different markets are important. Finally, the development of so-called "product platform strategies" has shown that technology and demand may jointly shape firm diversification. There are of course limits to this diversification, which is why the modal number of HS6 products exported by firms is only small.[7] Demand or resource cannibalization effects and the tendency to move away from the core technological competence bound the number of products firms can efficiently produce and sell on the market. However many products the firm exports, goods are related by demand and supply conditions and should be considered jointly. These demand and supply conditions, together with the strategic environment, consumer tastes and other market-specific characteristics generate significant variation in the product-mixes exported to different destinations. Moreover, our results, which emphasize the central role of both technology and demand in shaping firm product-mix, shed new light on the recent evidence showing that changes to firm product structure contribute significantly to the reallocation of resources within firms towards their most efficient use, thus fostering aggregate revenue-based productivity growth (Bernard et al., 2010; Goldberg et al., 2010).

More broadly, our work here reminds us that outcome of strategic choice is the entire structure of the product mix, embedding both technology and demand factors, which ultimately shapes firm productivity and profitability performance.

The remainder of the paper is organized as follows. Section 2 describes the data used in the empirical analysis and documents the degree of export-sales variability across destinations.

---

[6]This interpretation requires that the external transfer of this excess capacity be subject to some form of market failure.

[7]Amongst firms exporting more than one product (multi-product firms), the modal number of products exported in both Italy and France is two, with median figures of seven and six, respectively.
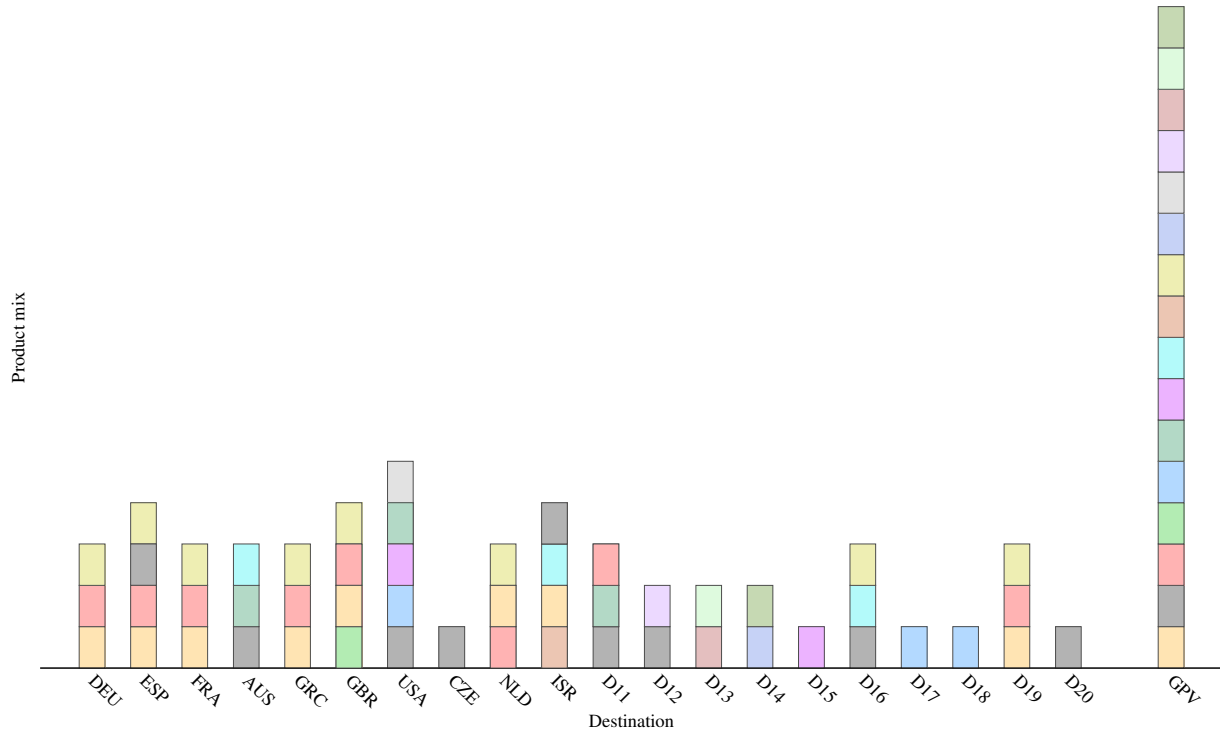
Figure 1: Local lists of products exported to each destination (ISO code) together with the global list of products exported worldwide (GPV). Products are ranked using market shares in each destination and globally respectively. Each color represents an HS6 code.

Section 2 defines the concept of firm product vectors, and Section 3 then investigates the degree of geographical variation in these vectors and discusses three major causes of this variation. Section 4 focuses instead on the stable part of exports. The last section concludes.

2   Data and definitions

Our analysis of multi-product and multi-country firms exploits trade data on the universe of French and Italian manufacturing exporters.

To illustrate what is new in our approach, we start with a real example taken from our data set. We consider a typical firm producing electrical motors, shipping 16 different products to 20 different destinations. We do not identify this firm and omit, on purpose, some information (e.g. on

certain destinations) for confidentiality reasons. This example is illustrated in Figure 1, where we show, by destination, the products exported, ranked by shares of local export value. Note that each color represents a different product and that the last "GPV" bar is the list of products exported by this firm ranked by share of worldwide export value. In each bar the product with the highest export value is at the bottom and that with the lowest export value at the top.

We first note that the combinations of products exported by this firm appear to a large extent fickle across markets: it exports 14 different product-mixes to the 20 active destinations. However, together with this variability we can also see that the firm exports some combinations of products more frequently than others.

Looking in more detail at this firm, it exports the same product mix to four out of its top 10 destinations[8] (DEU,FRA,GRC,NLD): "AC Motors" of two different powers plus an additional generic good labeled "Parts of these motors". In three of the other top-10 destinations (AUS, USA, CZE) sales are concentrated on a different product mix based on "Universal AC/DC motors". In addition, the relative importance of the three products exported to the top-EU destinations, while stable as a first approximation, changes across markets. Exports to Germany are mainly of high-power AC motors (of between 750W and 75KW), while in the Netherlands the greatest export share comes from low-power AC motors (with power below 750W).

Figure 1 provides illustrative evidence that goes against the usual idea that the export product ranking should be the same across destinations: the firm's worldwide (GPV) product hierarchy does not appear in any of the 20 destinations. In this example, the explanation is simple: the US is a large market, and the first product exported to the US accounts for a large share of total firm exports. And this product is specific to US demand patterns in terms of electrical motors.

We now proceed by describing the data set and define our novel approach to firm product

---

[8]In Figure 1 destinations on the X-axis are ranked according to their share of total export value. The top 10 destinations cover 90% of the firm's exports.

Table 1: DESCRIPTIVE STATISTICS

	Italy		France	
	Whole sample	Restricted sample	Whole sample	Restricted sample
	(1)	(2)	(3)	(4)
Total exports (billion Euros)	271.1	266.2	268.7	265.1
# Firms	74,365	45,530	32,432	19,689
Avg. # Products	8.4	12.6	8.5	12.9
Avg. # Countries	9.5	14.7	8.8	13.5

Notes: This table shows the descriptive statistics for the Italian and French dataset in 2006. In the restricted sample we keep those firms exporting more than one product and serving more than one destination.

mix. Our foreign-trade statistics are collected by the National Statistical Offices and consist of all the cross-border transactions of Italian and French firms.[9] The annual data on firm-product-destination level export flows are in values (Euros) and quantities (kilograms). Product categories are classified according to the Harmonized System and are reported at the six-digit (HS6) level.[10] As we are interested in the cross-section of firm-product exports across destinations, we restrict our attention to the most recent year available for both countries, namely 2006. Finally, since we focus on the variability of firm exports across products and countries, we exclude firms exporting only one product or serving only one destination.

Table 1 presents the total value of exports, the number of firms and the average number of products and destinations in the whole and restricted samples for Italy and France separately in 2006.[11] Note that while the number of firms falls substantially in the restricted sample is sizable

---

[9]The datasets were accessed at ISTAT and Banque de France facilities and were made available for analysis after careful screening to avoid disclosure of individual information.

[10]This is the finest level of disaggregation available for Italy, while the French customs data are available at the eight-digit (CN8) level. However, we consider the six-digit classification to make the analysis comparable across the two countries. This approach also gives us a better chance of finding the stable product lists the literature proposes.

[11]Table A1 in Appendix A reports the total value of exports and the number of exporters, distinguishing between three broad categories of firms: manufacturers, wholesalers and retailers, and others. The fraction of manufacturing exporters is larger in Italy than in France (55% vs 32%). These firms account for a large share of country exports (85% in Italy, 71% in France). On the contrary, France is characterized by a larger number of wholesale and retail exporters than Italy (43% against 36%). Wholesalers and retailers account for 21% of France's total exports but only 12% of those in Italy.
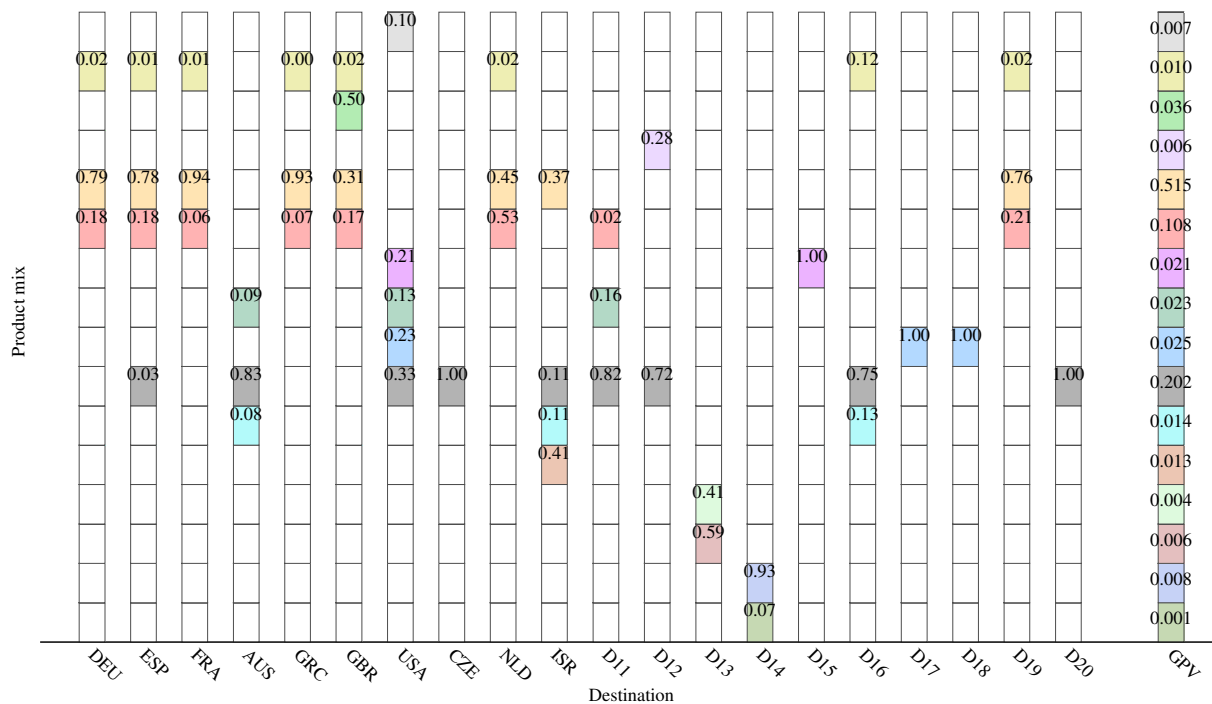
Figure 2: Local Product Vectors (LPV) exported to each destination, the Global Product Vector (GPV) exported worldwide and the corresponding market shares. To highlight the choice of not exporting an item, in each destination products are ranked according to the lexicographic order from the HS classification. Each color represents an HS6 code and "0.00" represents a market share under 1%.

in both countries, the value of exports changes only little.[12]

In the following, we characterize the export patterns of multi-country and multi-product firms. It is important to stress that we do not separately consider firm-destination and firm-product-destination sources of variation. We rather develop an empirical strategy, based on firm product mix heterogeneity across destinations, that to a large extent captures both sources of fickleness. We assess the extent of this variation and the factors behind it.

Product vectors

As opposed to previous empirical work that looked at each single good firms export we here focus on the mix of products that a firm sells in the different destinations where it is active. To this end we define the *Local Product Vector* (LPV_{fd}) and the *Global Product Vector* (GPV_f). The former is the set of products sold by firm f to destination d while the latter contains the whole set of products exported by firm f irrespective of their destination. In the remainder of this paper we analyze the LPVs and GPVs from two different but complementary perspectives.

The first, which is novel compared to the existing literature, is to look at product vectors as an ordered list of product names, in our case HS6 product codes. Formally, we represent the GPV_f as an ordered[13] vector of ones whose length is equal to the total number of products firm f exports (NP_f). Similarly, the LPV_{fd} vector is an ordered binary vector of ones when the product is actually exported by firm f to destination d and 0 otherwise. The LPV_{fd} is one of the 2NP_f possible arrangements of NP_f different products. As a result, all the local product vectors LPV_{fd} of a given firm have the same length as the corresponding GPV_f. This way of proceeding introduces a further novelty with respect to existing work in that it allows us to control for a firm's choice not to export a given product to a given destination: i.e. it takes into account product selection. The fact that a firm decides not to export one of its products to a given destination (when it does export it elsewhere) is an economic choice that likely contains relevant information.

Second, we add to the information in the LPVs and GPVs by including product market shares. Formally we replace the ones in the GPV_f by firm f's worldwide market share of the corresponding product and the ones in the LPV_{fd} by the market share of the corresponding product exported to destination d. By construction the GPVs and LPVs with market shares continue to have the

---

[12]As there are some missing values in the employment variable, in the analyses that follow we will work with a slightly smaller number of French firms (19,043 rather than 19,689).

[13]We follow the order of the HS classification system.

Table 2: EXPORT DIVERSIFICATION: PRODUCTS, DESTINATIONS and LPV

	Mean	Std. Dev.	Min.	1Q	2Q	3Q	Max.	Obs.
ITALY								
# Products	12.55	19.18	2.00	3.00	7.00	14.00	555.00	45,456
# Destinations	14.66	15.53	2.00	4.00	9.00	21.00	129.00	45,456
# LPV	8.46	9.80	1.00	3.00	5.00	10.00	117.00	45,456
# Products in LPV	2.43	2.37	1.00	1.33	1.80	2.62	65.31	45,456
FRANCE								
# Products	12.92	22.08	2.00	3.00	6.00	14.00	649.00	19,608
# Destinations	13.50	15.52	2.00	3.00	8.00	18.00	165.00	19,608
# LPV	7.98	9.90	1.00	2.00	4.00	9.00	141.00	19,608
# Products in LPV	2.59	3.00	1.00	1.33	1.80	2.72	75.73	19,608

Notes: This table shows the descriptive statistics of firm export diversification in 2006. The statistics are calculated on the restricted sample, i.e. after removing exporters with either a single product or a single destination.

same length, and in this case their components sum up to 1. Figure 2 shows the GPV_f and LPV_{fd} for the firm we used as an illustrative example above. Figure 2 further illustrates the richness of firm product mix. Adding to our remarks that the firm exports different product combinations to different destinations, and at the same time some combinations are more common than others, we note that, within each destination, the distribution of product sales is very skewed. Products with very high sales value are exported together with products with low sales value. However, the relative importance of each product differs by destination. Moreover, we can see that the non-export of a product to a destination is frequent, underlining once again the importance of controlling for product selection when looking at firm diversification in international markets.

Finally, moving away from the example, Table 2 presents the descriptive statistics for the product vectors of Italian (top panel) and French (bottom panel) exporters. The first two rows of each panel show the distribution of the number of products exported and destinations served by each firm. The average Italian exporter exports 12 products to more than 14 destinations, while French exporters export 12 products to 13 destinations.[14] Italian and French "average" exporters

---

[14]If we consider the whole population of exporters, the average number of products exported by both Italian and French firms is 8, and number of destinations is 9.5 for Italy and 8.8 for France, respectively.

are remarkably similar, in terms of both their product and geographical diversification. The third and the fourth row of each panel in Table 2 show, respectively, the descriptive statistics for the number of LPVs and of the number of products exported to a given destination (i.e. the number of ones in the LPV). The average number of LPVs in both Italy and France is 8 with, again on average, 2.5 products exported.[15] These two simple figures, beyond re-confirming the impressive similarity of the two countries, suggest that the average exporter does not sell the same product mix to each destination, and at the same time exports only a small subset of its products to each destination.

In the next section we investigate the degree of geographical heterogeneity in firm LPVs and discuss a number of the factors behind this diversity.

3 Product vectors: the fickle fringe

The quantitative assessment of the diversity of a firm's LPVs comes from, with a suitable metric, the distance between each LPV, first considered as a list of zeros and ones, and the corresponding GPV. The specific aim of this exercise is to see how common it is that a firm does not export a product to the whole set of destinations it serves.

We then change perspective and add the corresponding product market shares to the firms' GPVs and LPVs, calculating, again with a suitable metric, the distance of each LPV from its GPV. This second exercise is directly linked to the test of a perfectly rigid hierarchy of products in different destinations implied by the theoretical multi-product and multi-destination firm models mentioned above.

Finally, we carry out regressions to investigate the relative importance of three potential causes of the geographical diversification of Italian and French exporters.

---

[15]Note that a firm's maximum number of product vectors is the minimum of 2^{NP_f-1} and ND_f, where NP_f is the number of products exported and ND_f the number of destinations served by firm f.

Measuring diversity among product vectors

To capture the variability in a firm's local product vectors, as lists of zeros and ones, across destinations we measure the distance of LPV_{fd} from the corresponding GPV_f using a normalized version of the Levenshtein distance (LevD), also known as the Edit distance. This is a string metric developed to measure the difference between two sequences. Formally, it represents the minimum number of single edits (insertion, deletion, substitution) required to change one sequence into the other divided by the number of elements in the longest sequence minus 1.[16] Consider the following simple example for a firm exporting four different HS6 products to two destinations. Each element of the vector takes the value one if the corresponding product is exported, and zero otherwise. We thus consider the firm's binary choice to export a product or not.

$$
\text{GPV}_f = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{LPV}_{fd_1} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{LPV}_{fd_2} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} .
$$

Both LPV_{fd_1} and LPV_{fd_2} are different from the GPV_f, but how different are they? In this example $\text{LevD}(\text{GPV}_f, \text{LPV}_{fd_1})$ is 1/3: one change is needed to transform LPV_{fd_1} into the GPV_f and the number of elements in the longest sequence is four. Instead, the $\text{LevD}(\text{GPV}_f, \text{LPV}_{fd_2})$ is equal to 2/3: two changes are required to transform LPV_{fd_2} into the GPV_f and the length of the longest sequence is still four. More generally, in our example introducing an additional difference between the two sequences (by changing a one in a LPV to a zero) affects the Levenshtein distance by 1/3.

Figure 3 illustrates the frequency distribution of the LevD between each LPV_{fd} and the corre-

---

[16]See Appendix B for more details regarding this measure. Note that in our case we need to consider only substitutions since all the product vectors have the same dimension. In this case this metric is also known as Hamming distance.
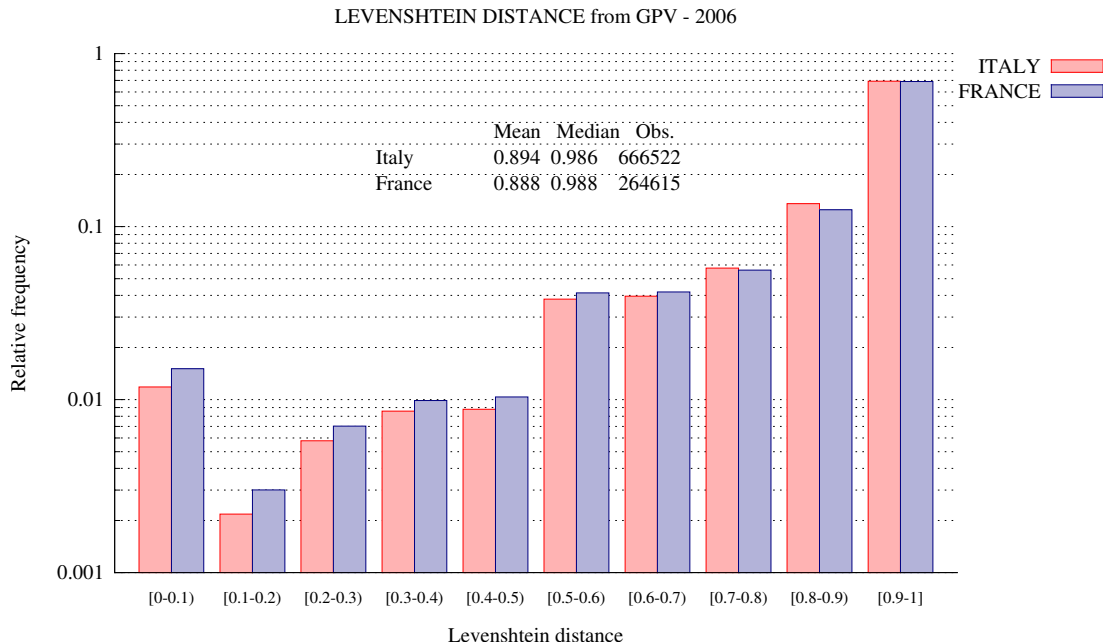
LEVENSHTEIN DISTANCE from GPV - 2006

	Mean	Median	Obs.
Italy	0.894	0.986	666522
France	0.888	0.988	264615

Figure 3: Frequency distributions of the Leveshtein distance from GPV in 2006 for Italy (red) and France (blue).

sponding GPV_f for Italian and French exporters. The observed values span the whole theoretical range of the Levenshtein distance, that is $[0, 1)$. However, it is apparent that the vast majority of the observations are concentrated in the right tail of the two distributions, with an average figure not far off of 0.9 and an even higher median value. Given that the average number of products of both Italian and French exporters is about 13, the average LevD figures means that we need to make around 11 changes to transform a local into the corresponding global product vector. This confirms that exporters in both countries have a wide export product portfolio, but export only a small and diverse subset of products to each destination.

We check whether this picture is driven by some particular sample properties via a number of robustness checks. We first ask whether firm size matters. We construct quartiles of the firm-size distribution in terms of number of employees and look at the LevD for the first and fourth quartiles. Second, we consider highly-diversified firms (those exporting over 15 products) versus

14

little-diversified firms (those exporting fewer than 6). Third, we carry out a stress test by calculating the LevD using only the firm's most important destination in term of export value. Last, we replicate the exercise using data from 2003. Appendix C, which reports the frequency distributions of the observed LevD for these robustness checks, largely confirms the picture given in Figure 3.

Working with product vectors as simple lists of HS6 codes has the principal advantage of allowing us to identify the effects of different forms of product complementarity (demand- or supply-driven) that can generate product combinations with equal parts of high- and low-value items: an example is Apple Inc., which exports both its I-pad and the dedicated Apple pencil.

However, it is clear that not all the products a firm exports to a destination play the same strategic competitive role. This is an important concern, since we already know that the firm distribution of export sales across products within a destination tends to be highly skewed, with sales concentrated on few items. Existing work on multi-product exporters has so far almost exclusively focused on product mixes defined using product market shares, and calculated rank correlations (Mayer et al., 2014; Arkolakis and Muendler, 2010; Di Comite et al., 2014).

While it is certainly important to weight products by their market shares, we believe that the existing rank-correlation approach suffers from two main drawbacks. First, rank correlation does not control for the firm's decision of whether to export a product to a given destination or not, which may bias upward the observed correlation. Second, rank correlation fails to quantify correctly the compositional dissimilarity between two different vectors, as it is based on rankings. We address both of these issues in turn below.

We first calculate the Spearman rank correlation, directly controlling for the firm's decision not to export a product to a given destination. The results are substantially lower than those in the existing literature: the mean and the median rank correlation fall to 0.41 and 0.42 for both Italy

and France.[17] Controlling for the extensive-margin decision then produces a substantial departure from a stable hierarchy of products across destinations.[18]

Second, we replace Spearman rank correlations by a distance measure to capture the effective dissimilarity between product vectors measured in terms of market shares. Rank correlation measures cannot capture the type of heterogeneity across PVs in a situation such as that illustrated below:

$$
\text{GPV}_f = \begin{pmatrix} 0.97 \\ 0.02 \\ 0.01 \end{pmatrix} \quad \text{LPV}_{fd_1} = \begin{pmatrix} 0.51 \\ 0.49 \\ 0 \end{pmatrix} \quad \text{LPV}_{fd_2} = \begin{pmatrix} 0.85 \\ 0.15 \\ 0 \end{pmatrix}
$$

Although LPV_{fd_1} and LPV_{fd_2} are very different, the rank correlation between each LPV and the GPV is one. We overcome this limitation via the Bray-Curtis dissimilarity measure of the compositional diversity between two vectors. This dissimilarity index is defined as

$$
BC_{fd} = \frac{\sum_p |s_{fp} - s_{fdp}|}{\sum_p |s_{fp} + s_{fdp}|} \quad , \tag{1}
$$

where s_{fp} and s_{fdp} represent the export shares of product p in the global vector and the destination vector d of firm f, respectively.[19] $\sum_p |s_{fp} + s_{fdp}|$ is a normalizing factor, which always equals two in our case. The Bray-Curtis dissimilarity measure is bound between 0 and 1, where 0 means the two vectors have the same composition ($s_{fp} = s_{fdp}$ for all p) and 1 means the two vectors are completely disjoint. In the example above the BC dissimilarity figures for LPV_{fd_1} and LPV_{fd_2} are

---

[17]Without considering the zeros, i.e. the choice of not exporting a product to an active destination, the mean and median Spearman pairwise rank correlations for Italy, using 2006 data, are 0.58 and 0.90 respectively. Similar figures are observed for France.

[18]The sharp reduction in rank correlation holds even when we restrict our analysis to the set of the most relevant markets served by a firm: those representing more than 1% of a firm's total exports. The figures from this robustness check are available upon request.

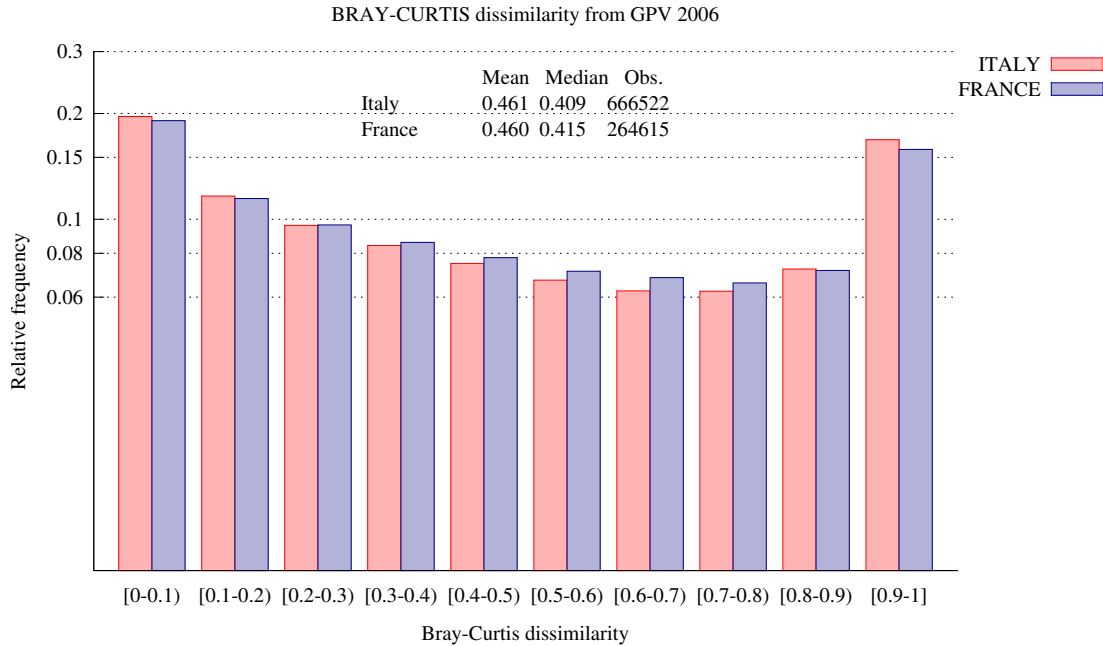[19]See Appendix B for more details regarding this measure.

Figure 4: Distribution of the Bray-Curtis distance from product-market shares of the GPV in 2006 for Italy (red) and France (blue).

0.47 and 0.13 respectively.

Figure 4 shows the Bray-Curtis dissimilarity measure between each firm LPV and the corresponding GPV. It is first of interest to note that, as expected, focusing on product-level export shares *reduces* the overall degree of heterogeneity between firm LPVs with respect to their GPV. However, only one fifth of observations have a BC value under 0.1, in both Italy and France. The average BC value of around 0.46 provides further evidence that there is not a very stable hierarchy among the products firms export to different destinations.

As before, we test the robustness of this investigation disaggregating our sample by firm size, the overall number of products exported, focusing only on the top destination, and using data from 2003. The results, in Appendix C, overall confirm the ability of the BC distance to capture heterogeneity across PVs and the robustness of our results.

Regression analysis

The analysis so far has revealed substantial geographical diversity among the export product vectors of Italian and French firms, and an associated departure from a perfect ordering of the products sold in different destinations. At the same time, the observed pattern is also at odds with products being randomly assigned across countries. In this section, we carry out regressions to identify the different factors behind the firm's choice of product mix exported to a given destination market.

Our point of departure is that theoretical models of multi-product multi-country firm export behavior based solely on cost efficiency cannot account for all the observed geographical variation in product mix, even though firm-product cost heterogeneity is an empirically-important determinant of firm export behavior. Our results are instead in line with models predicting an only imperfect correlation between the products a firm exports across countries, as in Bernard et al. (2011). Following this literature, we focus on three potential explanations.

The first of these is *Market size*. Demand heterogeneity, reflecting asymmetric preferences across varieties between countries, likely helps explain why firm product market share and ranking differ by destination. In this respect, the model in Di Comite et al. (2014) shows that single-product firms' sales vary across markets when each destination has different preferences for the same variety.

The second is *Market structure*, with the underlying idea that the competition the firm faces in a destination market likely influences the products exported there. In Mayer et al. (2014), for example, a firm responds to tougher competition by dropping its worst-performing products and concentrating its export sales on its best performers. In their model, however, changes in competition do not affect the rigid ordering of a firm's product ranking, but only product mix skewness.

The last explanation, which is not explicitly considered in any theoretical model, is the cross-country difference in firm market positioning with respect to its competitors. Firm export choice

can be related to the price of the product exported to a specific country relative to the average product-destination price of its competitors. This price difference can be driven by a number of factors, such as quality differences, markup heterogeneity, market competition, firm composition, supply factors such as shipping costs, and other destination-country characteristics. Regardless of the source of the difference in price within-firm across destinations, the firm's price relative to its competitors can explain the variation in firm-product export sales across countries. We refer to this idea as *Market positioning.*

We construct empirical counterparts for Market size, Market structure and Market positioning using trade data from BACI, a product-level dataset of imports and exports for a very large set of countries (see Gaulier and Zignago, 2010).

Our market-size measure is $IMP_{pd}^* = \sum_{o \in O_{pd}} IMP_{pod}^*$, the total imports of product p by destination d. Here O_{pd} is the set of countries exporting product p to destination d, excluding in turn Italy or France.[20] We then calculate the simple average[21] of $(\log)IMP_{pd}^*$ for each firm-destination pair as

$$\text{Market size}_{fd} = \frac{1}{|\Pi_f|} \sum_{p \in \Pi_f} \log IMP_{pd}^*$$

where $|\Pi_f|$ is the cardinality of the set of products exported by firm f. Higher values of Market size$_{fd}$ reflect greater potential demand for the set of products exported by firm f to destination d.

Next, our Market structure measure is $HHI_{pd}^* = \sum_{o \in O_{pd}} (Imp_{pdo}^* / \sum_{o \in O_{pd}} Imp_{pdo}^*)^2$, the Herfindal index for product p in destination d. We take the average over products for each firm-destination pair:

$$\text{Market structure}_{fd} = \frac{1}{|\Pi_f|} \sum_{p \in \Pi_f} \log HHI_{pd}^* \ .$$

---

[20] All of the measures from BACI data are identified by a $*$ superscript.

[21] A weighted average, with the relative importance of product p in firm total export sales as the weight, would introduce an endogeneity problem, as product market shares are also used in the construction of one of our dependent variables, the Bray-Curtis measure.

A higher value of Market structure$_{fd}$ shows more market concentration faced by firm f in foreign market d.

Last, we construct an empirical counterpart of Market positioning in a two-step procedure. We first calculate $UV^*_{pd} = \exp\left(\sum_{o \in O_{pd}} \left(\frac{IMP^*_{pod}}{IMP^*_{pd}} \right) \log UV^*_{pod} \right)$, the weighted[22] geometric average of the unit values of product p in destination d. Second we calculate $-\log |UV_{fpd} - UV^*_{pd}|$ the negative of the absolute value of the difference between the unit value of product p charged by firm f in destination d from the average in that market. Last, as above we take a simple average across the products exported by firm f to destination d:

$$\text{Market positioning}_{fd} = \frac{1}{|\Pi_{fd}|} \sum_{p \in \Pi_{fd}} \left(-\log |UV_{fpd} - UV^*_{pd}| \right) \quad .$$

The higher the value of this index, the closer is the market positioning of firm f in terms of unit values to those of its competitors.

We evaluate the empirical importance of these three factors in explaining LPV heterogeneity via the following regression

$$Y_{fd} = \beta_1 \text{Market size}_{fd} + \beta_2 \text{Market structure}_{fd} + \beta_3 \text{Market positioning}_{fd} + \alpha_f + \alpha_d + \epsilon_{fd} \quad , \quad (2)$$

where Y_{fd} is either the Levenshtein distance (LevD$_{fd}$) or the Bray-Curtis dissimilarity index (BC$_{fd}$). As we are modeling the dissimilarity index for a given firm, we include firm fixed effects α_f in the regression to account for any systematic ability differences between exporters that might affect trade outcomes across destinations. We also include destination fixed effects, α_d, which implicitly account for cross-country differences in total income, market toughness, trade costs and other institutional frictions affecting the firm's local product vector. We are interested in the β coef-

---

[22]The weights capture the relative importance of that transaction in the total imports of the product-destination pair pd.

Table 3: REGRESSIONS OF THE LEVENSHTEIN AND BRAY-CURTIS DISTANCE FROM GPV, 2006

	ITALY		FRANCE	
	(1a)	(1b)^a	(2a)	(2b)^a
Dependent variable: Lev$_{fd}$				
Market size$_{fd}$	-0.025	-0.266	-0.036	-0.411
	(0.002) [0.000]	(0.025) [0.000]	(0.004) [0.000]	(0.042) [0.000]
Market structure$_{fd}$	0.031	0.057	0.036	0.066
	(0.007) [0.000]	(0.014) [0.000]	(0.008) [0.000]	(0.016) [0.000]
Market positioning$_{fd}$	-0.002	-0.023	-0.001	-0.012
	(0.0004) [0.000]	(0.004) [0.000]	(0.0005) [0.050]	(0.006) [0.050]
Adj. R^2	0.54		0.55	
Dependent variable: BC$_{fd}$				
Market size$_{fd}$	-0.031	-0.173	-0.034	-0.221
	(0.002) [0.000]	(0.013) [0.000]	(0.002) [0.000]	(0.015) [0.000]
Market structure$_{fd}$	0.047	0.046	0.040	0.042
	(0.005) [0.000]	(0.005) [0.000]	(0.006) [0.000]	(0.007) [0.000]
Market positioning$_{fd}$	-0.023	-0.145	-0.014	-0.094
	(0.002) [0.000]	(0.010) [0.000]	(0.001) [0.000]	(0.006) [0.000]
Adj. R^2	0.49		0.48	
No. Obs	664,342		228,312	
Firm FE and Country FE	Yes		Yes	

Notes: This table shows Levenshtein and Bray-Curtis dissimilarity measure regressions. The data are from 2006. ^{a} Columns (1b) and (2b) show the standardized coefficients. Standard errors are clustered at the firm and country level. The coefficients appear together with standard errors (in brackets) and p-values (in square brackets).

ficients, which show the conditional correlation between the distance or dissimilarity index and our firm-destination level variables. As the right-hand side of the regression includes the logarithm of a geometric mean, the estimated coefficients can be interpreted as semi-elasticities. The error term, ϵ_{fd}, includes other firm-destination idiosyncratic factors that explain the observed heterogeneity. Standard errors are clustered at the firm and country level.

Table 3 contains the results for both the dependent variables. Columns (1a) and (2a) present the benchmark specification for Italy and France, respectively, while columns (1b) and (2b) list the beta coefficients on the standardized variables. As our dependent variables are measured using a

scale that does not have a clear quantitative economic interpretation, and since we want to compare the importance of the different regressors, our discussion focuses on these latter results. Once again, we interpret the sign of β's as the sign of a conditional correlation that does not reflect causality.

The top panel shows the results from Levenshtein distance LevD_{fd} regressions. First, all the three firm-country level variables are significantly correlated with the distance measure. Second, in line with economic intuition, the dissimilarity between LPVs and the corresponding GPV is smaller in larger markets, in countries where there is less concentration and when the firm's match in terms of market positioning is better. Firms are more likely to export all the products in their portfolio to destinations where there is a greater demand for its products, where competition is not impeded by just a few large players, and where the unit values charged are in line with prevailing average unit values. Third, Market size seems to have the strongest effect. A one standard-deviation rise in market size reduces the Levenshtein distance by 0.266 and 0.411 standard deviations in Italy and France, respectively: this corresponds to a fall of one in the average number of changes required to transform a firm's LPV into its GPV for the typical exporter in both Italy and France. Market structure and market positioning have smaller effects. Last, we again see an impressive degree of similarity between Italian and French exporters, which suggests that there are some common mechanisms at work here (see the Discussion section below).

The bottom panel of Table 3 contains the results for the Bray-Curtis index. The analysis here is not only in terms of the variation in the composition of the product vectors but also the relative size of the export shares of each product in the vector. The results are very much in line with those for the LevD distance measure. Larger market size reduces the dissimilarity between the export shares of the products in the firm's local vector and those in the global vector. Firms operating in destinations with a greater demand for their products are more likely to export a vector of products similar to their global export vector. The variable Market structure is instead positively correlated

with the dissimilarity index: where concentration is higher, firms export product vectors with product market shares that are more dissimilar to those in the GPV. Last, firms export product vectors with shares that are more similar to their global vector to destinations where their product prices are in line with those of their competitors. In terms of relative importance, market size again plays the major role, although market positioning is more important here than in the top panel of the table. This is especially true in Italy, where firms are more sensitive to their relative fit in terms of the unit values in the destination market.

Robustness checks

This section presents a set of exercises that test the robustness of our results with respect to a number of confounding factors.

First, as suggested by a recent empirical literature on multi-product exporters, the firm's export product mix in different markets may be affected by tariffs. Bernard et al. (2011) show, for example, that falling trade costs cause firms to drop their least-attractive products, while Dhingra (2013) finds that export-oriented firms in Thailand reduce their product lines in response to a unilateral tariff cut.[23] We thus add an additional regressor, Tariff_{fd}[24] to specification (2), defined as

$$\text{Tariff}_{fd} = \frac{1}{|\Pi_f|} \sum_{p \in \Pi_f} \text{Tariff rate}_{pd}$$

where Tariff rate$_{pd}$ is the tariff paid to export product p to destination d, and $|\Pi_f|$ is the cardinality of the set products exported by firm f. We expect that higher tariffs faced by firm f in destination d will make it more difficult for the firm to export the whole set of products in its GPV, producing greater dissimilarity between LPVs and the GPV.

---

[23]Qiu and Zhou (2013) propose a theoretical model where efficient firms expand their export product scope in response to foreign tariff cuts, whereas those of inefficient firms fall.

[24]The variable Tariff is taken from the WITS (World Bank) database. We use the Most Favored Nation.

Table 4: REGRESSIONS OF THE LEVENSHTEIN AND BRAY-CURTIS DISTANCES FROM GPV: ROBUST-
NESS CHECKS FOR ITALY

	ITALY				
	Tariff	EU countries	Developed	No carry along	No MNC
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Lev_{fd}					
Market size$_{fd}$	-0.267	-0.126	-0.237	-0.245	-0.260
	(0.026) [0.000]	(0.020) [0.000]	(0.025) [0.000]	(0.022) [0.000]	(0.025) [0.000]
Market structure$_{fd}$	0.055	0.027	0.053	0.055	0.055
	(0.014) [0.000]	(0.009) [0.000]	(0.014) [0.000]	(0.013) [0.000]	(0.014) [0.000]
Market positioning$_{fd}$	-0.023	-0.022	-0.021	-0.019	-0.024
	(0.005) [0.000]	(0.005) [0.000]	(0.005) [0.000]	(0.004) [0.000]	(0.004) [0.000]
Tariff$_{fd}$	0.005				
	(0.005) [0.397]				
Adj. R^2	0.25	0.66	0.58	0.56	0.54
Dependent variable: BC_{fd}					
Market size$_{fd}$	-0.174	-0.107	-0.153	-0.129	-0.169
	(0.013) [0.001]	(0.034) [0.001]	(0.023) [0.000]	(0.012) [0.000]	(0.013) [0.000]
Market structure$_{fd}$	0.045	0.016	0.045	0.042	0.045
	(0.005) [0.000]	(0.006) [0.006]	(0.007) [0.000]	(0.006) [0.000]	(0.005) [0.000]
Market positioning$_{fd}$	-0.144	-0.055	-0.120	-0.138	-0.141
	(0.010) [0.000]	(0.008) [0.000]	(0.013) [0.000]	(0.010) [0.000]	(0.010) [0.000]
Tariff$_{fd}$	-0.002				
	(0.004) [0.512]				
Adj. R^2	0.49	0.56	0.51	0.45	0.49
No. Obs	654,375	279,289	451,856	652,562	618,387
Firm FE	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes

Notes: This table shows regressions of the Levenshtein and Bray-Curtis dissimilarity measures. The data are from 2006. The standard errors are clustered at the firm and country level. The standardized coefficients are shown together with standard errors (in brackets) and p-values (in square brackets).

Table 5: REGRESSIONS ON THE LEVENSHTEIN AND BRAY-CURTIS DISTANCE FROM
GPV: ROBUSTNESS CHECKS FOR FRANCE

	FRANCE			
	Tariff	EU countries	Developed	No carry along
	(1)	(2)	(3)	(4)
Dependent variable: Lev_{fd}				
Market size$_{fd}$	-0.421	-0.173	-0.341	-0.364
	(0.042) [0.000]	(0.018) [0.000]	(0.031) [0.000]	(0.034) [0.000]
Market structure$_{fd}$	0.069	0.032	0.057	0.065
	(0.016) [0.000]	(0.006) [0.000]	(0.017) [0.000]	(0.015) [0.000]
Market positioning$_{fd}$	-0.013	-0.037	-0.006	-0.009
	(0.006) [0.048]	(0.005) [0.000]	(0.007) [0.380]	(0.005) [0.110]
Tariff$_{fd}$	0.010			
	(0.004) [0.034]			
Adj. R^2	0.31	0.60	0.44	0.34
Dependent variable: BC_{fd}				
Market size$_{fd}$	-0.226	-0.139	-0.194	-0.132
	(0.015) [0.000]	(0.017) [0.000]	(0.017) [0.000]	(0.016) [0.000]
Market structure$_{fd}$	0.043	0.021	0.046	0.040
	(0.007) [0.000]	(0.004) [0.000]	(0.007) [0.000]	(0.007) [0.000]
Market positioning$_{fd}$	-0.092	-0.056	-0.085	-0.092
	(0.006) [0.000]	(0.007) [0.000]	(0.008) [0.000]	(0.006) [0.000]
Tariff$_{fd}$	-0.003			
	(0.002) [0.272]			
Adj. R^2	0.48	0.56	0.49	0.42
No. Obs	222,907	101,432	156,813	217,502
Firm FE	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes

Notes: The table shows regression of the Levenshtein and Bray-Curtis dissimilarity measures. The data are from 2006. The standard errors are clustered at the firm and country level. The standardized coefficients appear together with standard errors (in brackets) and p-values (in square brackets).

The estimated beta coefficients on Tariff$_{fd}$ appear in column 1 of Table 4 for Italy and Table 5 for France. Tariffs do not explain our results. Firms export product vectors closer to their global vector in larger markets and where their position in terms of export prices is more in line with that of their competitors. On the contrary, concentration is associated with greater distance. The Tariff$_{fd}$ variable is statistically insignificant in the regressions for Italy and barely significant for France in the Levenshtein regression. The insignificance of Tariff$_{fd}$ may raise doubts about the validity of this robustness check, so we also estimate our regression only on those products that are exported to EU countries, where there are no tariffs: the results appear in column 2 of both tables. As expected, all the coefficients are smaller here than in the whole sample, but the overall story remains the same.

Second, to check that our empirical results do not reflect product-mix heterogeneity caused by destination-country income, we re-calculate the distance measures and re-run the regressions focusing only on a set of developed countries.[25] A number of empirical contributions have shown that firms are more likely to export high-quality and technologically-advanced products to high-income countries (Verhoogen, 2008; Crinò and Epifani, 2012; Flach and Janeba, 2013). More generally, firms segment markets and adapt product quality and price according to destination-country income. The results, in column 3 of both tables, confirm that our results also hold in exports to developed countries only. The three firm-country level variables continue to explain export product-mix vectors across destinations.

Third, column 4 shows the results from a fourth sensitivity check regarding "carry-along trade", in which manufacturing firms export products that they do not produce (Bernard et al., 2012). In principle, we would need information on both production and exports at the product level to identify carry-along firms. As these data are not available, we make an approximation by excluding

---

[25]We define developed countries as those with *per capita* income levels above the 50th percentile according to the World Bank.

products that are contemporaneously exported and imported by the same firm. The findings are robust to this change in product vector composition, excluding the possibility that our results reflect the import of products which are then re-exported by the firm and (possibly) not produced by it.

Fourth, we remove multinational companies (MNCs), which are complex organizations that sprawl across industries and countries. As data on multinationals are available only for Italy, this robustness check only appears in column 5 of Table 4.[26] Our baseline results regarding the role of all three firm-country specific variables in explaining the distance between the firm's local and global product vector continue to hold.

Additional robustness checks appear in Table D1 for Italy and Table D2 for France in Appendix D. We ask whether our results change if we modify the way the three firm-destination level regressors are defined. We above constructed these by averaging over all the products a firm exported to a destination. This has the advantage of avoiding the endogeneity problems that arise in weighted averages but, at the same time, does not take into account the relative importance of each product in a destination. We account for this by re-calculating the three measures using only the top product in each destination for each firm. Second, we apply weights given by the total exports of all Italian firms of product p to destination d over the total exports of all Italian firms to destination d, excluding the exports of the firm under consideration. We also check that our findings are not a statistical artifact produced by the structure of the product classification by considering those firms that export over more than one HS2-digit product.[27] Finally, we replicate the analysis using 2003 data and using wholesalers instead of manufacturing firms. Our results are robust to all of these tests.

---

[26]Among the work on the behavior of multi-product multinationals see the recent contribution of Yeaple (2013).

[27]These latter firms represent, both in Italy and France, around 80% of the population, suggesting a remarkable degree of diversification among firms.
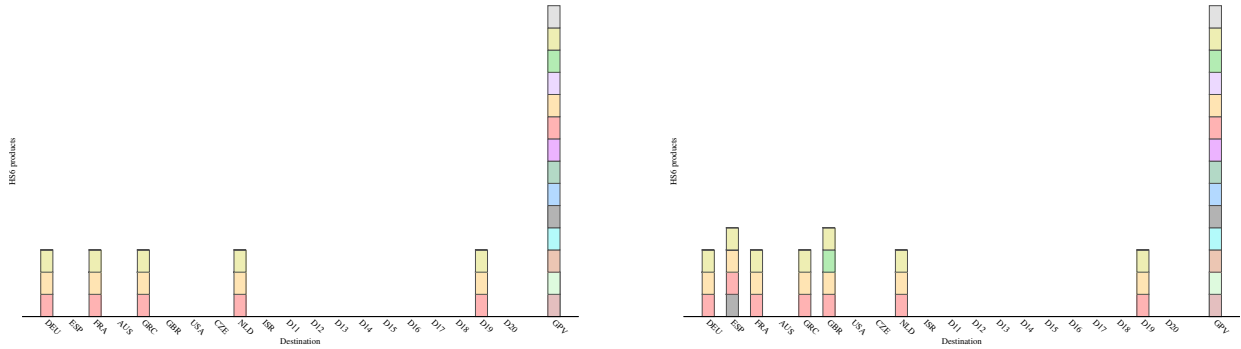
Figure 5: Core Product Vectors (CPV) and Weak Core Product Vectors (WCPV) for a firm in the Italian sample.

4 Product vectors: the stable core

The analysis of product vectors revels substantial heterogeneity in firm product mix across destinations, both from a qualitative and a quantitative point of view. At the same time, however, we do not see complete differentiation in LPVs, suggesting that there might be a structural component in the mix of products a firm exports that is stable across destinations. So far the literature on multi-product firms has focused on the concept of core products, defined as the top-selling varieties in which a firm seems to have some systematic advantage. Beyond considering single goods rather than vectors of products, existing work has also concentrated exclusively on the importance of a product in terms of export share. However, as in the illustrative example in the introduction, there are some combinations of products that are more likely to be exported than others. We have also demonstrated that there may be goods that, although less relevant in a purely quantitative sense, are crucial for the definition of a firm's typical product mix. Product complementarity or technological relatedness may imply that the production and thus export of one product leads to the production and export of its components or of other related goods that are not necessarily characterized by high sales.

We introduce two new concepts to provide a statistical characterization of this stable part of

27

Table 6: WCPV DESCRIPTIVE STATISTICS

	Mean	Std. Dev.	Min.	1Q	2Q	3Q	Max.	Obs.
ITALY - Whole sample								
Number of Products in WCPV	8.36	14.98	1.00	2.00	4.00	9.00	385.00	45,456
Share of destinations with WCPV	0.62	0.26	0.01	0.44	0.60	0.85	1.00	45,456
Share of exports covered by products in WCPV	0.80	0.31	0.00	0.70	0.97	1.00	1.00	45,456
FRANCE - Whole sample								
Number of Products in WCPV	8.78	18.64	1.00	2.00	3.00	8.00	586.00	19,608
Share of destinations with WCPV	0.62	0.26	0.02	0.46	0.62	0.85	1.00	19,608
Share of exports covered by products in WCPV	0.79	0.31	0.00	0.67	0.97	1.00	1.00	19,608

Notes: This table shows the descriptive statistics for firms' WCPVs in 2006. The statistics are calculated on the restricted sample, i.e. we remove exporters with either a single product or a single destination from the dataset.

product vectors: the *Core Product Vector* (CPV_f) and the *Weak Core Product Vector* ($WCPV_f$). The former is the most frequent arrangement observed across markets for firm f, and is measured by the typical (modal) product vector exported by firm f.[28] The WCPV is any arrangement which leaves all the cells with 1 in the CPV_f unchanged. In other words, the $WCPV_f$ is a product vector containing the CPV_f. Figure 4 shows the CPV_f and the $WCPV_f$ for the firm we use as an illustrative example here. The most frequent combination of products this firm exports is composed of three goods. Two of these are specific types of electric motors, and account for 52% and 11% of firm exports. The third is labeled "parts of motors", with a market share of only 1%. This combination of products is served in 5 out of 20 destinations. We also observe the very same product mix together with an additional idiosyncratic product in two other countries. Therefore 7 out of 20 destinations are served with the WCPV. It is also of interest to note that six out of the seven are in the firm's Top-10 destinations in terms of market shares, while the other one accounts for only 1%.

Table 6 presents the descriptive statistics for firms' WCPV in Italy (top panel) and France (bottom panel) regarding the number of products in the WCPV, the share of destinations covered

---

[28]Regarding the identification of the CPV_f, in case of ties the LPV_f are ranked by their value. In other words they are ranked according to the total value exported to destination d.

Table 7: THE PROBABILITY OF EXPORTING THE WCPV, 2006.

	ITALY		FRANCE	
	(1a)	(1b)[a]	(2a)	(2b)[a]
Dependent variable: DWCPV_{fd}				
Market size$_{fd}$	0.023	0.091	0.027	0.127
	(0.003) [0.000]	(0.012) [0.000]	(0.003) [0.000]	(0.015) [0.000]
Market structure$_{fd}$	-0.039	-0.027	-0.035	-0.027
	(0.005) [0.000]	(0.004) [0.000]	(0.008) [0.000]	(0.006) [0.000]
Market positioning$_{fd}$	0.013	0.056	0.012	0.061
	(0.001) [0.000]	(0.006) [0.000]	(0.001) [0.000]	(0.007) [0.000]
Adj. R^2	0.32		0.38	
No. Obs	664,342		228,312	
Firm FE	Yes		Yes	
Country FE	Yes		Yes	

Notes: This table shows WCPV regression results. The data are for 2006. [a] Columns (1b) and (2b) show standardized coefficients. Standard errors are clustered at the firm and country level. The coefficients are listed together with standard errors (in brackets) and p-values (in square brackets).

by WCPVs, and the share of exports covered by WCPV products. We first note that the average number of unique products in the whole set of WCPVs is about eight in both Italy and France. This is substantially fewer than 13, the average number of products exported in the two countries. Second, a WCPV is exported, on average, to about 60% of the destinations served by the firm, again in both Italy and France. Third, WCPV products account for a large share of firm total exports (on average about 80%). Last, the average export share of products in a WCPV is around 10%.

Overall these statistics suggest the existence of a set of products that is a strict subset of all the goods the firm exports and that is sold in a good number of its active destinations. The sales distribution within this subset is very skewed, but overall this set of products appears to be substantial in terms of its part of total exports. For these reasons we identify these as the firm's stable core of its product portfolio.

In the remainder of this section we investigate the factors behind the choice of exporting a WCPV to a given destination, focusing, in line with the previous analyses, on the role of market size, market structure and market positioning. We thus estimate the following regression

$$DWCPV_{fd} = \beta_1 \text{Market size}_{fd} + \beta_2 \text{Market structure}_{fd} + \beta_3 \text{Market positioning}_{fd} + \alpha_f + \alpha_d + \epsilon_{fd} \ ,$$

(3)

where DWCPV_{fd} is a dummy for firm f exporting the WCPV to destination d, $\text{X}_{fd,k}$ are our three firm-destination specific regressors, and α_f and α_d are firm and destination fixed effects, respectively. The explanatory variables are logarithms of geometric means, so that we can interpret the estimated coefficients as semi-elasticities but, as previously, we will focus our discussion on the estimates from the standardized variables.

Columns (1b) and (2b) of Table 7 present the results. The picture that emerges is in line with the previous results and confirms the great degree of similarity between France and Italy. Market size is here also the most important factor. The coefficients on market size are positive and statistically significant: the greater is destination demand, the more likely the firm will export a product combination including its core product vector there. The estimated coefficients on market structure indicate that firms facing higher concentration are less likely to export their weak core product vectors. That is, firms tend to serve destinations with tougher competition with their core products. This result is in line with one of the findings in Mayer et al. (2014) that firms respond to greater competition by exporting the products that are most directly related to their core competencies. Last, the coefficients on market positioning are positive and statistically significant, so that firms with positioning that is closer to that of their market competitors are more likely to export their core products.

5 Discussion

The above empirical analysis has produced three facts regarding firm product mixes across destinations. First, using the novel concept of firm product vectors, we show that firm product mixes exhibit a great deal of variation across destinations both in terms of the combination of products and export value shares. Second, taste heterogeneity and market structure are the key firm-destination idiosyncratic factors explaining this variability. Third, firm product mixes have both a fickle part and a stable component across destinations: firms typically export some combinations of products to a number of different destinations, so that the product vector has a core part.

External validity

Concerns may be raised about the external validity of our results: Can we infer systematic patterns in terms of stability and the variance of exporters' product mix in other countries from them? Even with the usual caveats, our answer to this question is Yes. A first simple reason is that we have already shown that the results are very similar across two different countries, France and Italy.

Additional evidence of the external validity of our results comes from the existing empirical literature on multi-product and multi-destination exporters. Arkolakis and Muendler (2013, 2010) report three robust stylized facts concerning the number of products exporters ship (*export scope*) and the corresponding average sales per product (*export scale*) to each destination, which characterize firms in Brazil, Denmark, Chile and Norway. First, a few large wide-scope exporters and many narrow-scope firms coexist in each destination. Second, the sales of wide-scope exporters are concentrated on a few products, and the same firms are able to cope with lower sales for their lowest-selling products. Third, there is a systematic positive relationship between average exporter scale (i.e the average sales per product within each country) and exporter scope. To further check that there is nothing particular about our data, we replicate their analysis and find the same styl-

ized facts for both Italian and French exporters: see Appendix E. This similarity in the results gives us confidence that there is nothing particular about French and Italian exporters that drives our results.

Back to theory

The omnipresence and importance of multi-product firms in international trade has opened new avenues for theory. What do our results imply in light of the theoretical models that have been proposed to describe the production of multiple goods in international markets?[29]

While they differ in their assumptions regarding supply, all of these models make significant contributions to the literature and replicate some important stylized facts regarding multi-product exporters. These models are simplifying representations of the complex strategies of exporters Regarding our results here, all of the models adopt at least one of the following two simplifications. First, they do not directly focus on the variation in firm product mix across destinations, and often imply an almost perfectly-stable hierarchy of products across destinations. Second, they treat each product within the firm in isolation, without taking into account any inter-dependencies resulting from technology or demand factors.

The first piece of evidence we presented suggests, however, that even after controlling for firm-specific productivity and common destination attributes, firms' export product vectors vary significantly (the fickle fringe), with this variation being at least partly explained by firm-destination specific demand conditions, strategic environment and consumer tastes. As such, a successful model of multi-product firms in international trade should incorporate idiosyncratic firm attributes across products, and possibly also across export destinations as in Bernard et al. (2011). In this respect, our results call for a departure from the standard specifications of preferences used in trade models

---

[29]See Arkolakis and Muendler (2010), Eckel and Neary (2010), Bernard et al. (2011) and the recent contribution of Mayer et al. (2014).

in favor of an approach similar to that in Di Comite et al. (2014) in the context of single-product firms. In their theoretical framework the same product sold by a firm in different countries faces different demands depending on the interactions between local tastes and competition effects. Note again that while our results do indicate the importance of firm-country idiosyncratic tastes for firm export decisions, we do not claim that this is the only driver of the fickle fringe in product mixes: since we do not observe demand preferences directly, we cannot exclude alternative explanations for our findings.

At the same time we have shown that firm product vectors include a component that is common across destinations (the stable core). As such, models should also consider at least one source of interdependence across a subset of products within firms, a feature that does not yet appear in any existing international-trade models. In this respect promising insights may be found in the industrial-organization literature, which has extensively analyzed the causes of firm diversification and why certain combinations of products are more likely to be co-produced or co-demanded than are others. Among these, various types of economies of scope, strategic bundling/tying of products, incompatibility risks and "product platform strategies" are the most common.[30] The integration of these theoretical causes of diversification generating product complementarities into the standard models of multi-product exporting is a challenging but potentially far-reaching avenue for future research.

As a final remark, it is important to stress that extending multi-product export models along the lines sketched out above might also be of use for other relevant streams of research on multi-product firms emphasizing the importance of intra-firm adjustments and suggesting that these differ significantly from those that occur between firms due to entry and exit (Bernard et al.,

---

[30]Bayley and Friedlaender (1982) and Montgomery (1994) provide two classic reviews of the motives for diversification. Stigler (1963) and Whinston (1990) investigate product bundling/tying for price-discrimination and market-power leverage purposes, respectively. A typical example of demand complementarity associated with compatibility issues is in Yu and Wong (2015). Baldwin and Woodard (2009) provide an overview of production platforms. Roberts and McEvily (2005) and Silverman (1999) discuss the factors that limit diversification.

2010; Goldberg et al., 2010; Navarro, 2012; Soderbom and Weng, 2012). This research, showing that changes in firms' product mix have contributed significantly to both individual and aggregate output growth, has not addressed whether and how intra-firm reallocation can be affected by the interdependence of supply and demand across products.

6   Conclusion

Despite the recent growth in papers underlining the preeminence and role of multi-product firms in international trade, we only know little about the behavior of these firms on foreign markets. How do firms diversify their product portfolio across destinations? What are the main factors behind a firm's different product mix across markets?

While there are a number of theoretical models of multi-product multi-country firms in international markets, the existing theory does not take into account potential demand or technology interdependence across products that can drive firm product-mix choices. Our results here suggest that such production or demand complementarities make some combinations of products within a firm more likely to be exported than others.

Using information on the universe of Italian and French manufacturing exporters, we establish new stylized facts consistent with a more complex model where firms do not have a stable ranking of product mix across markets, but rather adapt their choices to better match destination characteristics. We use the concept of product vectors to analyze multi-product firm export behavior across destinations. By using sequences of product names we provide new insights on the extent to which a firm's product vectors change across destinations. We also employ a statistic that allows us to quantify the compositional diversity between two different vectors.

We first reveal a substantial degree of dissimilarity between a firm's local export product vector and its corresponding global vector, both from a qualitative and a quantitative perspective. Sec-

ond, we offer a number of potential explanations of this diversity. Market demand heterogeneity, reflecting asymmetric preferences for different varieties, market competition and market positioning emerge as significant factors in this respect.

Third, we confirm the existence of a stable set of products exported to different destinations but, as opposed to previous empirical work, we find that core product vectors include products that are relatively marginal in terms of export share. The goods in core product vectors are not necessarily characterized by high sales due to product complementarity or technological relatedness.

Overall, our empirical findings paint a more complex picture of the behavior of multi-product firms in foreign markets. Given the significance of these firms in international trade, we believe that it is important to understand their export patterns. One important avenue for future research is to understand how the firm's choice of products across destinations in our empirical analysis affects the welfare and distributional consequences of international trade.

References

ARKOLAKIS, C. AND M.-A. MUENDLER (2010): "The Extensive Margin of Exporting Products: A Firm-level Analysis," NBER Working Paper 16641, National Bureau of Economic Research, Inc.

——— (2013): "Exporters and their products: a collection of empirical regularities," *CESifo Economic Studies*, 59, 223–248.

BALDWIN, C. Y. AND C. J. WOODARD (2009): "The Architecture of Platforms: A Unified View," in *Platforms, Markets and Innovation*, ed. by A. Gawer, Edward Elgar.

BAYLEY, E. E. AND A. F. FRIEDLAENDER (1982): "Market structure and multiproduct industries," *Journal of Economic Literature*, 20, 1024–1048.

BERNARD, A. B., E. J. BLANCHARD, I. V. BEVEREN, AND H. Y. VANDENBUSSCHE (2012): "Carry-Along Trade," NBER Working Papers 18246, National Bureau of Economic Research, Inc.

BERNARD, A. B., J. B. JENSEN, S. J. REDDING, AND P. K. SCHOTT (2009): "The Margins of US Trade," *American Economic Review*, 99, 487–93.

BERNARD, A. B., S. J. REDDING, AND P. K. SCHOTT (2010): "Multiple-Product Firms and Product Switching," *American Economic Review*, 100, 70–97.

——— (2011): "Multi-Product Firms and Trade Liberalization," *Quarterly Journal of Economics*, forthcoming.

BRAY, J. R. AND J. T. CURTIS (1957): "An ordination of upland forest communities of southern Wisconsin," *Ecological Monographs*, 27, 325–349.

CHANEY, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–21.

CRINÒ, R. AND P. EPIFANI (2012): "Productivity, Quality and Export Behaviour," *Economic Journal*, 122, 1206–1243.

DHINGRA, S. (2013): "Trading Away Wide Brands for Cheap Brands," *American Economic Review*, 103, 2554–84.

DI COMITE, F., J.-F. THISSE, AND H. VANDENBUSSCHE (2014): "Verti-zontal Differentiation in Monopolistic Competition," *Journal of International Economics*, forthcoming.

EATON, J., S. KORTUM, AND F. KRAMARZ (2011): "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 79, 1453–1498.

ECKEL, C. AND J. P. NEARY (2010): "Multi-Product Firms and Flexible Manufacturing in the Global Economy," *Review of Economic Studies*, 77, 188–217.

FEENSTRA, R. AND H. MA (2008): "Optimal Choice of Product Scope for Multiproduct Firms under Monopolistic Competition," in *The Organization of Firms in a Global Economy*, ed. by E. Helpman, D. Marin, and T. Verdier, Harvard University Press.

FLACH, L. AND E. JANEBA (2013): "Income Inequality and Export Prices across Countries," CESifo Working Paper 4298, CESifo Group Munich.

FREUND, C. AND M. D. PIEROLA (2015): "Export superstars," *Review of Economics and Statistics*, 97, 1023–1032.

GAULIER, G. AND S. ZIGNAGO (2010): "BACI: International Trade Database at the Product-Level. The 1994-2007 Version," Working Papers 2010-23, CEPII research center.

GOLDBERG, P. K., A. K. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2010): "Multiproduct Firms and Product Turnover in the Developing World: Evidence from India," *Review of Economics and Statistics*, 92, 1042–1049.

KEE, H. L. AND K. KRISHNA (2008): "Firm-Level Heterogeneous Productivity and Demand Shocks: Evidence from Bangladesh," *American Economic Review*, 98, 457–62.

MAYER, T., M. J. MELITZ, AND G. I. P. OTTAVIANO (2014): "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, 104, 495–536.

MONTGOMERY, C. A. (1994): "Corporate diverification," *Journal of Economic Perspectives*, 8, 163–178.

NAVARRO, L. (2012): "Plant level evidence on product mix changes in Chilean manufacturing," *The Journal of International Trade & Economic Development*, 21, 165–195.

NOCKE, V. AND S. YEAPLE (2014): "Globalization And Multiproduct Firms," *International Economic Review*, 55, 993–1018.

QIU, L. D. AND W. ZHOU (2013): "Multiproduct firms and scope adjustment in globalization," *Journal of International Economics*, 91, 142–153.

ROBERTS, P. W. AND S. MCEVILY (2005): "Product-line expansion and resource cannibalization," *Journal of Economic Behavior and Organization*, 57, 49–70.

SILVERMAN, B. S. (1999): "Technological resources and the direction of corporate diversification: toward an integration of the resource-based view and transaction cost economics," *Management Science*, 45, 1109–1124.

SODERBOM, M. AND Q. WENG (2012): "Multi-product firms, product mix changes and upgrading: Evidence from China's state-owned forest areas," *China Economic Review*, 23, 801–818.

STIGLER, G. J. (1963): "United States v. Loew's Inc.: A Note on Block-Booking," *The Supreme Court Review*, 1963, 152–157.

VERHOOGEN, E. (2008): "Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 123, 489–530.

WHINSTON, M. D. (1990): "Tying, Foreclosure, and Exclusion," *American Economic Review*, 80, 837–59.

YEAPLE, S. R. (2013): "Scale, Scope, and the International Expansion Strategies of Multiproduct Firms," Working Paper 19166, National Bureau of Economic Research.

YU, C. AND T. N. WONG (2015): "A product bundle determination model for multi-product supplier selection," *Journal of Intelligent Manufacturing*, 26, 369–385.

Appendix A

Table A1 shows the total value of exports and the number of exporters, distinguishing between three broad categories of firms (manufacturers, wholesalers and retailers, and others) in France and Italy.

Table A1: Exports and Number of exporting firms: share by type of firms, 2006

	Manufacturers (1)	Wholesalers & Retailers (2)	Others (3)	Total (4)
ITALY				
Total exports (billion Euros)	271.1	38.5	9.1	318.7
# Firms	74,365	48,643	14,861	134,579
FRANCE				
Total exports (billion Euros)	268.7	79.8	28.3	376.9
# Firms	32,432	44,379	25,473	102,284

Appendix B

Levenshtein distance

The Levenshtein distance between two sequences is the minimum number of single edits (insertion, deletion, substitution) required to change one sequence into the other, divided by the number of elements of the longest sequence. Mathematically the Levenshtein distance between two sequences $s = \{s_1, \ldots, s_{N_s}\}$ and $q = \{q_1, \ldots, s_{N_q}\}$ is

$$Lev(s,q) = \frac{\text{Edit}(N_s, N_q)}{\text{Max}(N_s, N_q) - 1} \quad ,$$

where $\text{Edit}(N_s, N_q)$ is given by the following recursion

$$\text{Edit}(i,j) = \begin{cases} \text{Max}(i,j) & \text{if } \text{Min}(i,j) = 0 \\ \\ \text{Min} \begin{cases} \text{Edit}(i-1,j) + 1 \\ \text{Edit}(i,j-1) + 1 \\ \text{Edit}(i-1,j-1) + \mathbf{1}(s_i \neq q_j) \end{cases} & \text{otherwise} \end{cases}$$

in which $\mathbf{1}(\cdot)$ is the indicator function.

Bray-Curtis dissimilarity index

The Bray-Curtis dissimilarity index, also known as the Sørensen index, is a well-known way of quantifying the difference between samples.[31] Formally, the Bray-Curtis dissimilarity index between vector i and vector j is defined as

$$BC_{i,j} = \frac{\sum_{k=1}^{K} |i_k - j_k|}{\sum_{k=1}^{K} |i_k + j_k|} \quad , \tag{4}$$

where i_k and j_k represent the number of elements observed in vector i and j along the $k\text{th}$ dimension. The Bray-Curtis index is symmetric and ranges from 0, when the two vectors are identical, to 1 where the two vectors are disjoint.[32] The BC measure is not an Euclidean distance since it does not satisfy the triangular inequality axiom.

Usually, the BC index is computed on count data.[33] Consider the two following vectors i and j with dimension $K = 3$

$$i = \begin{pmatrix} 11 \\ 0 \\ 7 \end{pmatrix} \quad j = \begin{pmatrix} 24 \\ 37 \\ 5 \end{pmatrix} \quad ;$$

---

[31] Originally this measure was developed to study species abundance in different locations in ecological analysis.(Bray and Curtis, 1957)

[32] Two vectors i and j are disjoint if whenever there is a non-zero entry in i, there is a zero entry in j and vice versa.

[33] For example, the number of firms in a location or the number of products exported to a destination.

the $BC_{i,j}$ between the two is 0.619, and is obtained as follows:

$$BC_{i,j} = \frac{|11 - 24| + |0 - 37| + |7 - 5|}{18 + 66} = 0.619 \quad .$$

(5)

When calculated with raw count data, the BC dissimilarity index captures differences associated with both the *size* and the *shape* of the two vectors, where the former refers to differences in the total number of elements, $\sum_k i_k$ and $\sum_k j_k$, in vector i and j respectively while *shape* concerns the distribution of elements along the different dimensions of the two vectors.

We here calculate the BC dissimilarity index between a firm's GPV and its LPVs where the elements of the vectors are product export shares. In this case, the denominator in equation (4) is always 2, and the index captures differences only in *shape*. The BC index is particularly useful when we observe firms exporting products to different destinations with the same ranking but with very different export shares. Consider the simple case in which a firm's export share vectors are

$$\text{GPV}_f = \begin{pmatrix} 0.98 \\ 0.02 \end{pmatrix} \quad \text{LPV}_{f1} = \begin{pmatrix} 0.99 \\ 0.01 \end{pmatrix} \quad \text{LPV}_{f2} = \begin{pmatrix} 0.51 \\ 0.49 \end{pmatrix} \quad .$$

It is straightforward to see that the rank correlations between GPV and the two LPVs are both 1 and that they provide an only imprecise picture of the diversity of local vectors relative to the global vector.

Appendix C

Figure C1: Distribution of the Leveshtein distance from GPV for different size classes (**Top-left panel**), for firms exporting fewer than six products (**Top-right panel**) and more than 15 products (**Center-left panel**), using only the most important destination in terms of export value for each firm (**Center-right panel**) and in 2003 (**Bottom-left panel**).

Figure C2: Distribution of the Bray-Curtis distance from GPV for different size classes (**Top-left panel**), for firms exporting fewer than six products (**Top-right panel**) and more than 15 products (**Center-left panel**), using only the most important destination in terms of export value for each firm (**Center-right panel**) and in 2003 (**Bottom-left panel**).

Appendix D

The weighted versions of the three firm-level measures are defined as

$$\text{Market size}_{fd} = \frac{1}{|\Pi_f|} \sum_{p \in \Pi_f} W_{fp} \left(\log IMP^*_{pd} \right)$$

$$\text{Market structure}_{fd} = \frac{1}{|\Pi_f|} \sum_{p \in \Pi_f} W_{fp} \left(\log HHI^*_{pd} \right)$$

$$\text{Market positioning}_{fd} = \frac{1}{|\Pi_{fd}|} \sum_{p \in \Pi_{fd}} W_{fpd} \left(-\log |UV_{fpd} - UV^*_{pd}| \right) \quad .$$

To limit the risk of introducing potential endogeneity, the weights are defined as

$$W_{fp} = \frac{EXP_p - EXP_{fp}}{\sum_f (EXP_p - EXP_{fp})}$$

$$W_{fpd} = \frac{EXP_{pd} - EXP_{fpd}}{\sum_f (EXP_{pd} - EXP_{fpd})} \quad ,$$

where EXP_p is the export value of product p exported by all the firms in Italy or France and EXP_{fp} is the export value of product p for firm f.

Table D1: REGRESSIONS OF THE LEVENSHTEIN AND BRAY-CURTIS DISTANCE FROM GPV, 2006: ADDITIONAL ROBUSTNESS CHECKS FOR ITALY

| | ITALY | | | | |
| | Top products | Weights | HS2 | Year 2003 | Wholesalers |
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Lev_{fd}					
Market size$_{fd}$	-0.047	-0.195	-0.268	-0.276	-0.278
	(0.007) [0.000]	(0.023) [0.000]	(0.024) [0.000]	(0.024) [0.000]	(0.025) [0.000]
Market structure$_{fd}$	0.004	0.039	0.062	0.032	0.041
	(0.005) [0.389]	(0.012) [0.001]	(0.014) [0.000]	(0.012) [0.010]	(0.017) [0.014]
Market positioning$_{fd}$	-0.062	-0.015	-0.018	-0.018	-0.034
	(0.008) [0.000]	(0.004) [0.000]	(0.004) [0.000]	(0.005) [0.000]	(0.006) [0.000]
Adj. R^2	0.24	0.24	0.24	0.24	0.23
Dependent variable: BC_{fd}					
Market size$_{fd}$	-0.090	-0.155	-0.180	-0.182	-0.159
	(0.005) [0.000]	(0.011) [0.000]	(0.014) [0.000]	(0.012) [0.000]	(0.016) [0.000]
Market structure$_{fd}$	0.005	0.032	0.048	0.038	0.039
	(0.004) [0.203]	(0.005) [0.000]	(0.005) [0.000]	(0.005) [0.000]	(0.006) [0.000]
Market positioning$_{fd}$	-0.164	-0.095	-0.145	-0.119	-0.095
	(0.011) [0.000]	(0.007) [0.000]	(0.010) [0.000]	(0.008) [0.000]	(0.010) [0.000]
Adj. R^2	0.50	0.49	0.50	0.48	0.45
No. Obs	664,114	666,518	595,644	643,337	150,500
Firm FE and Country FE			Yes		

Notes: This table shows the regression of the Levenshtein and Bray-Curtis dissimilarity measures. The data are for 2006. Column 1: the three regressors are calculated using only the most important product for the firm in that destination. Column 2: we use different weights. Column 3: only firms exporting over more than one HS2 sector. Column 4: we run the regressions for 2003. Standard errors are clustered at the firm and country level. Standardized coefficients are listed together with standard errors (in brackets) and p-values (in square brackets).

Table D2: REGRESSIONS OF THE LEVENSHTEIN AND BRAY-CURTIS DISTANCE FROM GPV, 2006: ADDITIONAL ROBUSTNESS CHECKS FOR FRANCE

	FRANCE				
	Top products	Weights	HS2	Year 2003	Wholesalers
	(1)	(2)	(3)	(4)	(5)
Dependent variable: Lev_{fd}					
Market size$_{fd}$	-0.046	-0.256	-0.423	-0.369	-0.381
	(0.006) [0.000]	(0.027) [0.000]	(0.045) [0.000]	(0.039) [0.000]	(0.017) [0.000]
Market structure$_{fd}$	-0.003	0.031	0.067	0.034	0.051
	(0.004) [0.538]	(0.015) [0.041]	(0.017) [0.000]	(0.017) [0.042]	(0.007) [0.024]
Market positioning$_{fd}$	-0.029	0.025	-0.013	-0.002	-0.014
	(0.006) [0.000]	(0.006) [0.000]	(0.006) [0.050]	(0.005) [0.634]	(0.006) [0.000]
Adj. R^2	0.26	0.26	0.29	0.26	0.31
Dependent variable: BC_{fd}					
Market size$_{fd}$	-0.102	-0.148	-0.239	-0.178	-0.257
	(0.008) [0.000]	(0.015) [0.000]	(0.017) [0.000]	(0.016) [0.000]	(0.023) [0.000]
Market structure$_{fd}$	0.003	0.028	0.047	0.051	0.044
	(0.003) [0.436]	(0.007) [0.000]	(0.007) [0.000]	(0.008) [0.000]	(0.010) [0.000]
Market positioning$_{fd}$	-0.111	-0.056	-0.095	-0.097	-0.029
	(0.007) [0.000]	(0.004) [0.000]	(0.006) [0.000]	(0.007) [0.000]	(0.008) [0.000]
Adj. R^2	0.52	0.48	0.48	0.48	0.45
No. Obs	205,522	264,615	212,337	263,744	69,507
Firm FE and Country FE			Yes		

Notes: This table shows the regression of the Levenshtein and Bray-Curtis dissimilarity measures. The data are for 2006. Column 1: the three regressors are calculated using only the most important product for the firm in that destination. Column 2: we use different weights. Column 3: only firms exporting over more than one HS2 sector. Column 4: we run the regressions for 2003. Standard errors are clustered at the firm and country level. Standardized coefficients are listed together with standard errors (in brackets) and p-values (in square brackets).

Appendix E

This appendix shows that the same regularities discussed in Arkolakis and Muendler (2013, 2010) characterize both Italian and French exporters. Figure E1 shows the relationship, for Italian firms (top panel) and French firms (bottom panel), between exporter scope and the corresponding percentile in the exporter-scope distribution for firms shipping to Germany and to the USA (similar patterns are found for other destinations). The distribution is very skewed: more than 70% of exporters export only three or fewer products, while only the top 10% ship more than 10 products. In Figure E2 we plot, for firms exporting 4, 8, 16, and 32 products, the average (across firms) export product value for products sharing the same rank against their rank within firm. These figures confirm first that in Italy and France wide-scope exporters are indeed much larger, in terms of sales, than narrow-scope exporters. But more importantly, wide-scope exporters are able to cope with lower sales of their lowest selling products. In this respect, regressing the exporter's lowest-ranked product's (log) sales against (log) exporter scope in a market conditioning on fixed effects for firm and destination produces an elasticity of -1.45 (0.007) and -1.56 (0.011) for Italy and France, respectively (Arkolakis and Muendler (2010) estimate an elasticity for Brazilian firms of 2.1). Finally, Figure E3 confirms for Italian and French firms that within-destination mean exporter scope and the corresponding mean exporter scale are positively related.
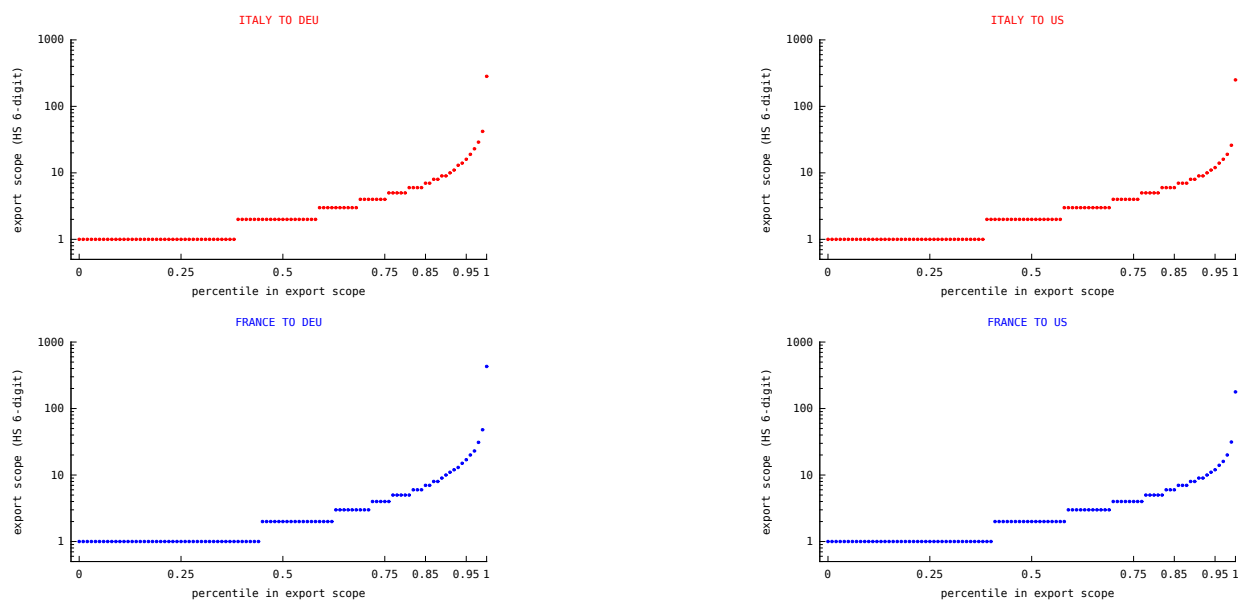

Figure E1: Italian (**Top panels**) and French (**Bottom panels**) exporter scope distribution regarding Germany (left panel) and the USA (right panel). Customs data from 2006. The products are at the H6 6-digit level. The scale of the vertical axis is logarithmic.
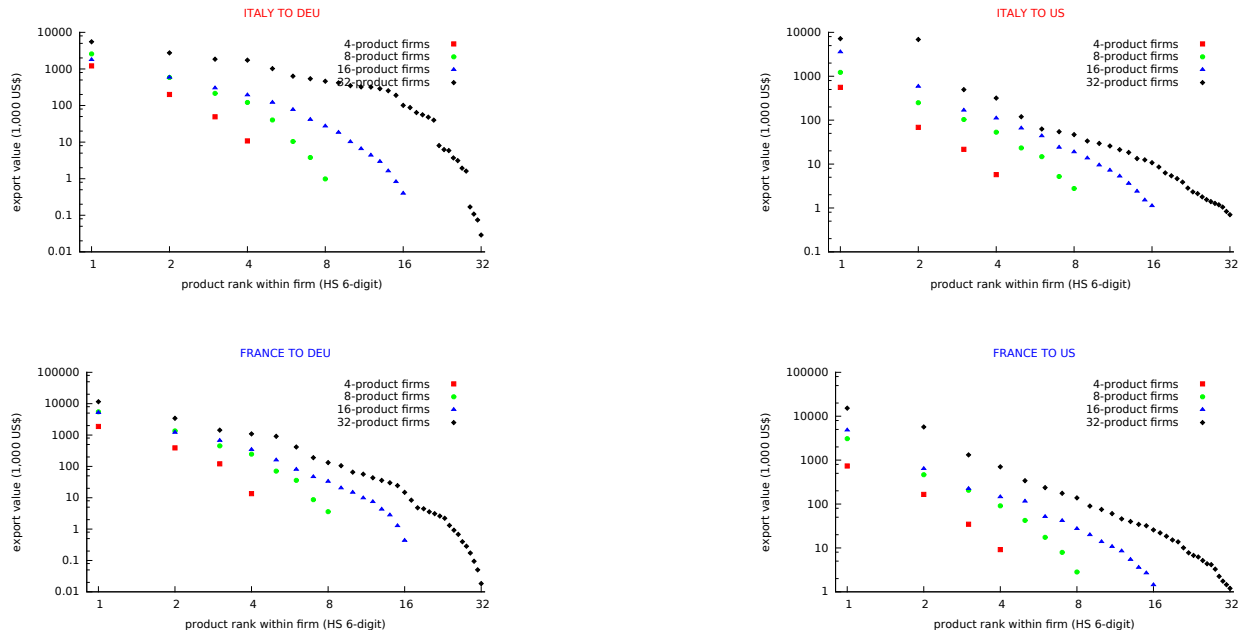
Figure E2: Italian (**Top panels**) and French (**Bottom panels**) firms' distributions of product sales regarding Germany (left panel) and the USA (right panel). The figure shows the distribution for firms exporting exactly 4, 8, 16 or 32 products. Customs data from 2006. The products are at the H6 6-digit level. The scales of both the vertical and horizontal axes are logarithmic.
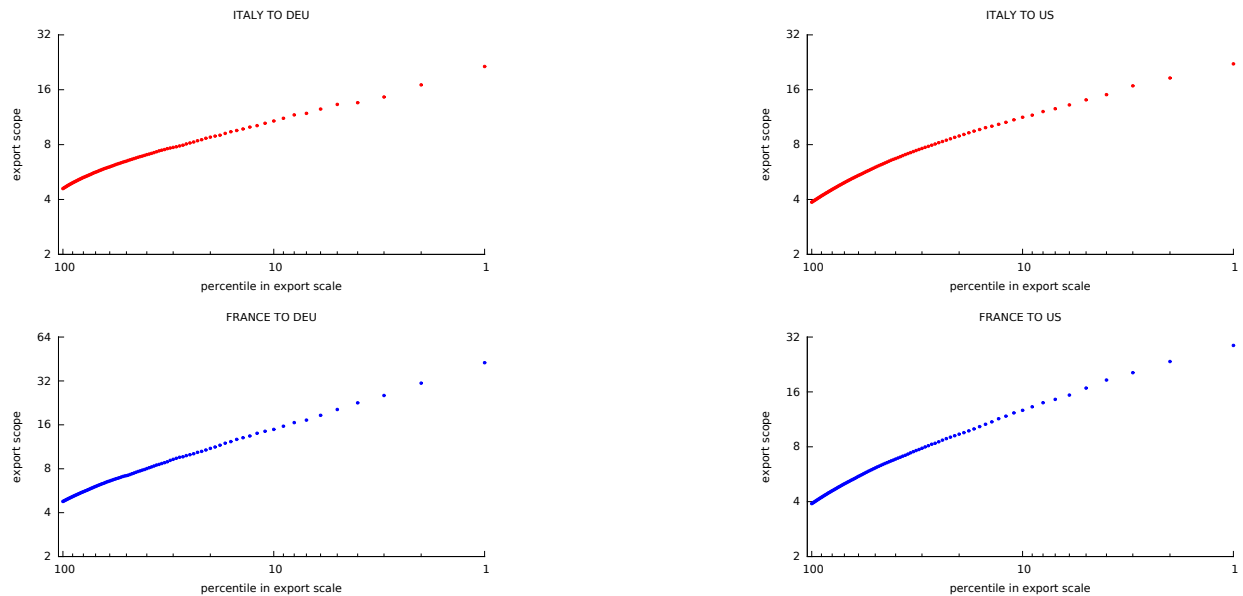

Figure E3: The relationship between mean Italian (**Top panels**) and French (**Bottom panels**) exporter scale and mean exporter scope regarding Germany (left panel) and the USA (right panel). Customs data from 2006. The products are at the H6 6-digit level. The scales of both the vertical and horizontal axes are logarithmic.

47