

Concentration in International Markets: Evidence from US Imports*

Alessandra Bonfiglioli[†] Rosario Crinò[‡] Gino Gancia[§]

This draft: February 2019

Abstract

We use transaction-level data to study changes in the concentration of US imports. Concentration has fallen in the typical industry, while it is stable by industry and country of origin. The fall in concentration is driven by the extensive margin: the number of exporting firm has grown, and the number of exported products has fallen more for top firms. Instead, average revenue per product of top firms has increased. At the industry level, top firms are converging, but top firms within country are diverging. These facts suggest that intensified competition in international markets coexists with growing concentration among national producers.

JEL Classification: E23, F12, F14, L11, R12.

Keywords: Superstar Firms, Concentration, US Imports, Firm Heterogeneity, International Trade.

*We thank Philippe Aghion, Timo Boppart and Peter Neary for their comments. We acknowledge financial support from the Spanish Ministry of Economy and Competitiveness (ECO2014-59805-P) and through the Severo Ochoa Programme for Centres of Excellence in R&D (SEV-2015-0563) and the Spanish State Research Agency MDM-2016-0684 under the Maria de Maeztu Unit of Excellence Programme.

[†]Queen Mary University of London and CEPR, Mile End Road, London E1 4NS, UK. E-mail: a.bonfiglioli@qmul.ac.uk

[‡]Università Cattolica del Sacro Cuore, Dept. of Economics and Finance, CEPR and CESifo. Via Necchi 5, 20123, Milan, ITALY. E-mail: rosario.cрино@unicatt.it.

[§]Queen Mary University of London, CREi and CEPR. Mile End Road, London E1 4NS, UK. E-mail: g.gancia@qmul.ac.uk

1 INTRODUCTION

We live in a superstar economy in which top firms command a disproportionate share of sales and wealth. A large number of papers have documented that, since the late 1990s, the fraction of sales accrued by top firms and other concentration indexes have risen in most US sectors. International evidence, albeit more sparse, indicates that concentration has grown in several OECD countries too. Large firms also dominate exports. In a sample of 32 mostly developing countries, the top five firms account on average for 30% of a country's total exports (Freund and Pierola, 2015). These observations have raised serious concerns that the growth of superstar firms may be synonymous of lower competition. The size of the phenomenon is so large that it is often the subject of media attention.¹ Yet, little is known to date on its causes and consequences.

Crucially, the existing evidence points at growing concentration among *national* firms. However, companies from different countries compete in markets that are increasingly global. This is especially true in the manufacturing sector, where import penetration is particularly high.² In global markets, stronger international competition and growing national concentration can coexist. In fact, leading models of international trade suggest that international competition causes reallocations towards top producers and may therefore increase concentration at the national level (e.g., Melitz, 2003, Melitz and Redding 2014).

In this paper, we examine concentration in *international markets*. To this end, we use a unique transaction-level data set to study changes in the concentration of US imports between 2002 and 2012. Focusing on imports allows us to complement the picture arising from national production data. It enables us to document how firms from multiple countries compete in the world's largest global market. We start by showing how concentration, measured by the share of total imports that is accrued by the four largest firms and the Herfindahl-Hirschman Index, has changed among firms selling to the United States in the average 4-digit manufacturing industry. We contrast this with changes in concentration among firms from a single country of origin and among US producers.

To get a sense of the geographical distribution of concentration, we draw maps showing

¹For instance, *The Economist* recently published a special report on the rise of giant companies (September 17, 2016). According to it, 10% of the world's public companies generate 80% of all profits, and the share of GDP generated by the Fortune 100 biggest American companies rose from about 33% of GDP in 1994 to 46% in 2013.

²Import penetration in US manufacturing is around 30%. It is significantly higher in some sectors, such as high-tech industries, and in smaller countries.

which countries have affected competition in the US import market the most. Next, we shed light on how concentration has changed. With the help of a simple structural model imposing minimal restrictions on the data, we dissect the observed variation in the sales share of the top four exporters into the contribution of the firm-level characteristics that we can extract from the data.

Our main findings are as follows. First, we show that while industry concentration among US firms has increased, concentration is stable among exporters from the average country of origin and has fallen significantly among all exporters. Regarding geographical patterns, industry concentration among exporters is lowest in Western Europe, South-East Asia, India and China, and has fallen especially in the latter two countries.

Second, the most important factor in explaining the fall in concentration is the extensive margin. We observe a large increase in the number of firms that start exporting to the US in any industry. The extensive margin plays an important role also within firms. While all firms are shedding products, top firms are dropping proportionally more products than other firms. On the contrary, the main force towards rising concentration is the intensive margin. Consistent with the view that top firms are pulling away, we show that the average revenue per product of top firms has grown significantly relative to the population average. The differential growth of top firm-products is stronger at the industry level than among firms from the same country of origin.

Third, of this differential growth, up to 30 percent is explained by divergence in performance within top firm-products at the industry level, and up to 20 percent at the industry-country level. In other words, revenue per product is increasingly more unequal both between top and non-top firms, and within the set of top firms. But top firms in any industry are more different than the rest and more equal to each other than top firms from the same country of origin. Finally, the results from our exact decomposition of top import shares are confirmed by regression analysis, where we can use alternative measures of concentration, add control variables, and explore different sources of variation in the data.

These findings paint a significantly different picture than the one arising from national-level studies. The sheer increase in the number of firm-products exported to the US suggests that the overall level of competition may have intensified rather than fallen, even if the number of US entering firms has declined. Growing global competition is also consistent with the observed within-firm adjustments. As the total number of products increases, products per firm are falling, suggesting that firms are concentrating on core business to retain a competitive edge. Top firms, which are likely to be present in more markets and hence more

exposed to global competition, are concentrating on their core business more than proportionally, instead of acquiring rapaciously new product lines. On the other hand, whatever the underlying reason might be, our results do show that firms are growing more and more unequal. Understanding the consequences of this process remains therefore a pressing question for future research.

This paper is related to the growing literature on the rise of national concentration. Several papers, including Gutierrez and Philippon (2017), Barkai (2017) and Autor, Dorn, Katz, Patterson, and Van Reenen (2017), have documented the recent increase in concentration among US firms. Other papers have tried to measure changes in competition more directly by estimating markups. In particular, De Loecker and Eeckhout (2017 and 2018), Hall (2018) and Diez, Leigh, and Tambunlertchai (2018) have found increases in average markups both in the US and other countries.³ However, this increase is partly explained, once again, by the growth of high-markup firms.⁴ Other studies, such as Karabarbounis and Neiman (2018) and Anderson, Rebelo, and Wong (2018), have found more mixed results. All these papers study concentration or markups among national firms using production data. Hence, they do not study how concentration changed in any market.

We are aware of only two exceptions. Rossi-Hansberg, Sarte and Trachter (2018) present evidence that US concentration, while growing at the national level, has actually been falling in local US markets. Their findings resonate well with ours. Nevertheless, they are entirely different in that Rossi-Hansberg, Sarte and Trachter (2018) still focus exclusively on US firms. While they correctly point out that many markets are local, making their analysis relevant, many others are global, especially in the manufacturing sector. Freund and Sidhu (2017) is instead a first attempt at measuring global industrial concentration.⁵ Interestingly, they also find that global concentration has declined in most industries, but they do so combining national data. Hence, they study concentration among producers in the world, not in any destination market.

Finally, this paper is related to the large empirical literature on the role of firms, and especially superstar firms, in explaining trade flows.⁶ Redding and Weinstein (2018) and

³See also Calligaris, Criscuolo and Marcolin (2018).

⁴Reallocations towards high-markup firms may actually imply lower monopoly distortions. See for instance Baqaee and Farhi (2017), Edmond, Midrigan and Xu (2018), and Epifani and Gancia (2011).

⁵Due to data limitations, only a few papers compare systematically firms across countries. See for instance, Gennaioli et al. (2013), Bartelsman, Haltiwanger and Scarpetta (2013), Bloom, Sadun and Van Reenen (2016) and Poschke (2015).

⁶See for instance Bernard, Jensen, Redding and Schott (2018), Freund and Pierola (2015), Gaubert and

Bonfiglioli, Crinò and Gancia (2018b) decompose US imports into firm-level characteristics. We build on their approach to document and decompose changes in concentration among firms exporting to the United States. Many papers have shown that trade opportunities trigger reallocations in favor of top firms, thereby making firms more unequal and possibly raising national concentration (see, for instance, di Giovanni, Levchenko and Ranciere, 2011). These reallocations can also happen within firms. In particular, Bernard, Redding and Schott (2011) and Melitz, Mayer and Ottaviano (2014) show that trade liberalization and tougher competition cause multi-product firms to drop their least successful products and skew export sales towards the best performing products. They provide supportive evidence using US and French firm-level data, respectively. Our findings on how international concentration has changed among firms from multiple origins are entirely consistent with the view proposed in these papers.⁷

The remainder of the paper is organized as follows. Section 2 provides a brief description of the data and reports key statistics on concentration measures and their evolution across firms from the same country of origin and industry, across all exporters to the United States in the same industry, and across US firms. In Section 3, we implement a structural decomposition of the variation in concentration of US imports as measured by the change in the sales share of top-four firms in each country-industry and industry. Section 4 studies by means of regressions how concentration correlates with the number of firms and average sales per firm, both at the country-industry and at the industry level. Section 5 concludes. In the Appendix, we provide more information on our data and details on the estimation of the elasticity of substitution needed to implement the structural decomposition.

2 TRENDS IN DOMESTIC AND IMPORT CONCENTRATION

2.1 DATA

To perform our analysis, we use transaction-level data on US imports from Piers, a database administered by IHS Markit. Piers contains the complete detail of the bill of lading of any container entering the US by sea. IHS Markit collects, verifies and standardizes the information contained in the bills of lading, and makes the resulting data available for sale. We purchased from IHS Markit information on the universe of waterborne import transactions

Itskhoki (2018).

⁷On the importance of the extensive margin in international trade, see Bonfiglioli, Crinò and Gancia (2018b), Fernandes, Klenow, Meleshchuk, Pierola and Rodriguez-Clare (2018) and Hummels and Klenow (2005).

of the US, by exporting firm and product, in two years, 2002 and 2012. For each transaction, we know the complete name of the exporting firm, its country of origin, the exported product (according to the 6-digit level of the HS classification), the value (in US dollars) and the quantity (in kilograms) of the transaction.⁸ We assign products to industries using a correspondence table between the HS 6-digit classification and the SIC 4-digit classification developed by the World Integrated Trade Solutions.

In our final data set, firms belong to 104 origin countries spanning the five continents, and sell products across 366 manufacturing industries. In the Appendix, we provide more details on the construction and composition of the data set, as well as on the data cleaning procedure we follow to exclude transactions potentially contaminated by reporting mistakes. The Appendix also present descriptive statistics about the coverage of Piers, in terms of both export value and number of firms. Although Piers registers waterborne transactions only, we find its coverage to be remarkably good, accounting for 83% of the total exports to the US for the average country.

2.2 CONCENTRATION STATISTICS

We measure industry concentration as the share of total sales that is accrued by the four largest firms in an industry and by the industry's Herfindahl-Hirschman Index.⁹ We compute these measures both by industry, pooling firms' sales from all origin countries, and by industry and country. In the latter case, we treat multinational firms selling from multiple countries as different firms, while we combine their sales when studying concentration at the industry level. For comparison, in this section we also report the corresponding concentration measures among US firms from COMPUSTAT.

Table 1 contains the main descriptive statistics. It shows that, in the average industry, the top-4 firms account for 79% of all imports by the average country (panel a) and 37% of imports from all countries (panel b). Interestingly, concentration among US firms is comparable to the level observed by country (panel c). The table also reports changes in the concentration indexes between 2012 and 2002. Over the decade, concentration among firms from the same country barely changed. However, concentration among firms from all countries decreased

⁸In the case of firms with multiple shipments (bills of lading) of the same product in a year, we purchased from IHS Markit information on the total value and quantity of these shipments across all bills of lading. We did not purchase the detailed information on each bill of lading, as this information would have been prohibitively expensive.

⁹Results are very similar when using alternative criteria, such as the share of the top-3 or top-5 firms.

Table 1: Descriptive Statistics on Concentration Measures

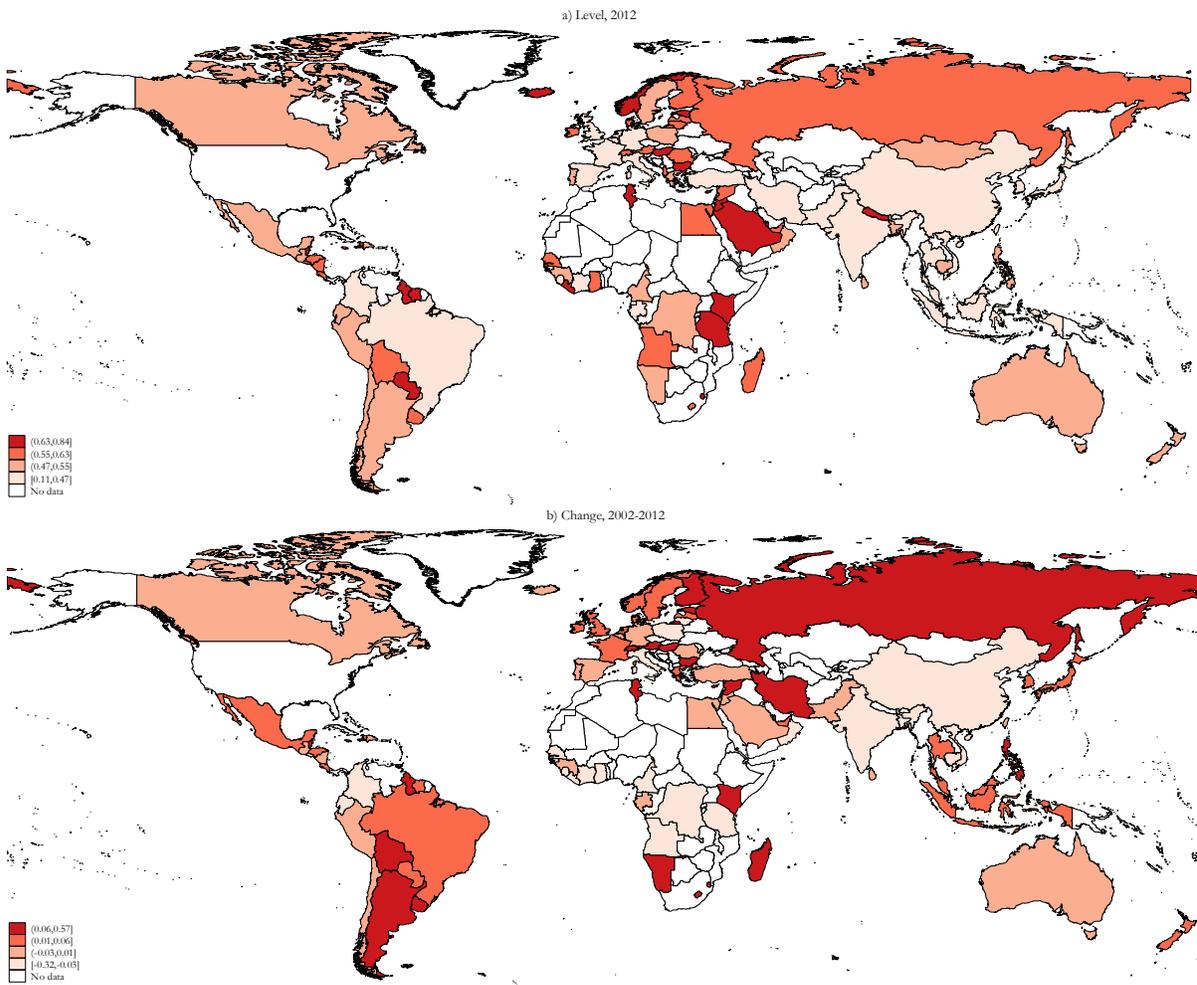
	Mean (2012)	Std. Dev. (2012)	Change (02-12)	% of Cases with Rise of Concentration
a) <u>PIERS: Statistics by country-industry pair</u>				
Share of sales by top-4 firms	0.79	0.21	-0.01	0.47
Herfindahl index	0.46	0.29	0.01	0.50
b) <u>PIERS: Statistics by industry</u>				
Share of sales by top-4 firms	0.37	0.23	-0.08	0.34
Herfindahl index	0.09	0.13	-0.03	0.32
c) <u>COMPUSTAT: Statistics by industry</u>				
Share of sales by top-4 firms	0.88	0.15	0.05	0.70
Herfindahl index	0.55	0.30	0.13	0.73

Notes. Industries are defined at the 4-digit level of the Standard Industrial Classification (SIC).

significantly: the share of sales by top-4 firms fell by 8 percentage points. Conversely, as it is well known, the top-4 share among US firms increased, by 5 percentage points. All these patterns hold when measuring concentration using the Herfindahl-Hirschman index, or when counting the fraction of industries becoming more concentrated over the period. These statistics suggest that rising concentration among national firms can coexist with more competition in international markets.

Our main goal is to dissect the changes in concentration in US imports just documented. As a preliminary step, to get a sense of the role played by individual countries, we take a look at the geographical distribution of concentration among firms selling to the United States. In Figure 1 we draw world maps showing the average level and the change in concentration by country. Darker colors indicate a higher level of the Herfindahl index in 2012 (map a) or a larger increase in the index between 2002 and 2012 (map b). All figures are based on country-level arithmetic averages computed across 4-digit industries.

Some geographical patterns stand out. Focusing on levels, concentration in 2012 appears to be lower than average in Western Europe, India, China and some parts of Southeast Asia. It is higher in parts of Eastern Europe, the Middle East and Russia. These cross-country patterns are broadly consistent with the evidence on the geographical distribution of markups reported in De Loecker and Eekhout (2018). Focusing on changes, concentration has grown in Latin America, Eastern Europe and Russia, and has fallen in China and India. These results are in line with Freund and Sidhu (2017), who stress the contribution of China and emerging-market countries to global competition using national data.



Source: Piers (IHS Markit), US import data for 2002 and 2012. Darker colors indicate a higher level of the Herfindahl index of concentration in 2012 (map a) or a larger increase in the index between 2002 and 2012 (map b). All figures are country-level arithmetic averages computed across 4-digit industries.

Figure 1: Concentration Across Countries

3 DECOMPOSING TOP FIRMS' SHARES IN US IMPORTS

We now derive a simple decomposition that allows us to quantify the contribution of various firm-level characteristics to the observed changes in concentration, as measured by top firms' shares. Building on Bonfiglioli, Crinò and Gancia (2018b), the characteristics that we can identify are the number of firms, products per firm, average sales per product and heterogeneity across firm-products. Since the decomposition can be applied to any subset of firms, we will use it to study changes in industrial concentration among foreign firms selling into the US market both from all countries and by country of origin. Given that our data covers seaborne trade, our decomposition will be implemented on the subset of firms exporting by sea.

3.1 A STRUCTURAL DECOMPOSITION

Consider an industry i composed of differentiated varieties. Preferences over these varieties are:

$$C(i) = \left\{ \sum_{\omega \in \Omega_i} [\gamma(\omega)c(\omega)]^{(\sigma_i-1)/\sigma_i} \right\}^{\sigma_i/(\sigma_i-1)}, \quad \sigma_i > 1, \quad (1)$$

where $c(\omega)$ is the quantity consumed of variety $\omega \in \Omega_i$, where Ω_i denotes the set of varieties available for consumption in industry i ; $\gamma(\omega)$ is a demand shifter sometimes interpreted as quality; and σ_i is the elasticity of substitution between varieties in industry i . Each variety is produced by a different firm; however, to make the model consistent with the data, a firm may produce more than one variety. Hence, ω refers to a firm-product pair. Denote by $p(\omega)$ the price of variety ω . Then, the minimum cost of one unit of the consumption basket $C(i)$ is given by the price index:

$$P(i) = \left[\sum_{\omega \in \Omega_i} \tilde{\gamma}(\omega)^{\sigma_i-1} \right]^{1/(1-\sigma_i)}, \quad (2)$$

where $\tilde{\gamma}(\omega) \equiv \gamma(\omega)/p(\omega)$ is a synthetic measure of "appeal" of variety ω . Revenue from sales of a variety with appeal $\tilde{\gamma}$ is:

$$r(\omega) = \tilde{\gamma}(\omega)^{\sigma_i-1} P(i)^{\sigma_i} C(i) = \frac{\tilde{\gamma}(\omega)^{\sigma_i-1}}{\sum_{\omega \in \Omega_i} \tilde{\gamma}(\omega)^{\sigma_i-1}} P(i) C(i). \quad (3)$$

This equation shows that the scale of $\tilde{\gamma}$ is irrelevant because what matters is its ratio to the sector aggregate.

We use this model to decompose the sales of top firms in US imports. Hence, $P(i)C(i)$ will be the value of US imports in industry i . We start by decomposing top shares by industry and country. Let $n^f(i, o)$ be the number of *firms* exporting to the United States from country o in industry i . Let $n^p(i, o)$ be number of *products per firm* and $\bar{r}(i, o)$ the *revenue per firm-product* from country o in industry i . Finally, we denote with a subscript *top* the corresponding numbers of firms, products per firms and revenue per firm-product of the top $X \in \mathbb{N}$ firms from country o in industry i . By these definitions, the expenditure share of the top X exporters among all exporting firms from origin o is:

$$s_{top}(i, o) \equiv \frac{n_{top}^f(i, o) \cdot n_{top}^p(i, o) \cdot \bar{r}_{top}(i, o)}{n^f(i, o) \cdot n^p(i, o) \cdot \bar{r}(i, o)}, \quad (4)$$

where the denominator is total sales in industry i from country o and the numerator is the corresponding sales by top firms only. $s_{top}(i, o)$ measures concentration among exporters from a given country and industry. Equation (4) can immediately be used to decompose changes of the top firms' import shares into an extensive margin (number of firms and products per firm) and an intensive margin (average revenue per firm-product):

$$\Delta \ln s_{top}(i, o) = -\Delta \ln n^f(i, o) + \Delta \ln n_{top}^p(i, o) - \Delta \ln n^p(i, o) + \Delta \ln \frac{\bar{r}_{top}(i, o)}{\bar{r}(i, o)}. \quad (5)$$

The advantage of the structural model is that it can be used to further decompose the intensive margin into firm-level characteristics, namely, the distribution of appeal. To this end, use (3) to write:

$$\frac{r(\omega)}{\bar{r}(i, o)} = \frac{\tilde{\gamma}(\omega)^{\sigma_i - 1}}{\mathbb{E}[\tilde{\gamma}(i, o)^{\sigma_i - 1}]}, \quad (6)$$

where $\mathbb{E}[\tilde{\gamma}(i, o)^{\sigma_i - 1}]$ is the arithmetic average of all $\tilde{\gamma}(\omega)^{\sigma_i - 1}$ sold from origin o in industry i . Normalizing $\mathbb{E}[\tilde{\gamma}(i, o)^{\sigma_i - 1}] = 1$, we can identify $\tilde{\gamma}(\omega)^{\sigma_i - 1}$ from deviations in sales from the country-industry average, $r(\omega) / \bar{r}(i, o)$. Finally, to retrieve the appeal of each firm-product, $\tilde{\gamma}(\omega)$, a value for the elasticity of substitution, σ_i , is needed:

$$\tilde{\gamma}(\omega) = [r(\omega) / \bar{r}(i, o)]^{1/(\sigma_i - 1)}. \quad (7)$$

We summarize in the Appendix how σ_i can be estimated from the data as in Redding and Weinstein (2017) or Bonfiglioli, Crinò and Gancia (2018b). Next, observe that:

$$\frac{\bar{r}_{top}(i, o)}{\bar{r}(i, o)} = \mathbb{E}[\tilde{\gamma}_{top}(i, o)^{\sigma_i - 1}].$$

Adding and subtracting $\{\mathbb{E} [\tilde{\gamma}_{top}(i, o)]\}^{\sigma_i-1}$ and taking differences:

$$\Delta \frac{\bar{r}_{top}(i, o)}{\bar{r}(i, o)} = \overbrace{\Delta \{\mathbb{E} [\tilde{\gamma}_{top}(i, o)]\}^{\sigma_i-1}}^{\Delta \text{ average appeal}} + \overbrace{\Delta \mathbb{E} [\tilde{\gamma}_{top}(i, o)^{\sigma_i-1} - \{\mathbb{E} [\tilde{\gamma}_{top}(i, o)]\}^{\sigma_i-1}]}^{\Delta \text{ heterogeneity in appeal}}. \quad (8)$$

Equation (8) decomposes the change in average revenues of top firm-products relative to the country-industry mean into two terms: the change in average appeal of top firm-products and the change in its dispersion. The first term measures whether all top firm-products are becoming on average better than the rest. The second term captures instead the role of differential growth in appeal within top firm-products: when $\sigma_i > 2$ sales are a convex function of $\tilde{\gamma}$ implying that more dispersion in appeal leads to higher average sales, because the increase in demand for better-than-average firm-products is greater than the decrease in demand for the remaining ones. Hence, this formula allows us to exactly quantify whether top firms are converging or diverging in their attributes and how this affects their total sales.

We can also apply these decompositions at the industry level, pooling firms from all origins:

$$\Delta \ln s_{top}(i) = -\Delta \ln n^f(i) + \Delta \ln n_{top}^p(i) - \Delta \ln n^p(i) + \Delta \ln \frac{\bar{r}_{top}(i)}{\bar{r}(i)}. \quad (9)$$

Imposing the normalization $\mathbb{E}[(\tilde{\gamma}(i)^{\sigma_i-1})] = 1$ and re-computing

$$\tilde{\gamma}(\omega) = [r(\omega) / \bar{r}(i)]^{1/(\sigma_i-1)} \quad (10)$$

we can write:

$$\Delta \frac{\bar{r}_{top}(i)}{\bar{r}(i)} = \Delta \{\mathbb{E} [\tilde{\gamma}_{top}(i)]\}^{\sigma_i-1} + \Delta \mathbb{E} [\tilde{\gamma}_{top}(i)^{\sigma_i-1} - \{\mathbb{E} [\tilde{\gamma}_{top}(i)]\}^{\sigma_i-1}]. \quad (11)$$

3.2 TOP IMPORT SHARES BY INDUSTRY AND COUNTRY

We now show the results of the decompositions in (5), (8), (9) and (11). As already noted, the decomposition of the intensive margin depends on the elasticity of substitution, σ_i . To check the sensitivity of the results to this parameter, we work with two alternative measures of σ_i , both estimated using our micro data and the model structural equations. More details on the estimation methods are reported in the Appendix. Henceforth, we will use the following notation to label the two elasticities: *reg. contr.* will denote the estimate obtained from the regressions in Bonfiglioli, Crinò and Gancia (2018b) and *RW* the reverse-weighting estimate

Table 2: Decomposition of the Share of Sales by the Top-4 Firms

(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta \ln s_{top}(i, o)$	$-\Delta \ln n^f(i, o)$	$\Delta \ln n_{top}^p(i, o)$	$-\Delta \ln n^p(i, o)$	$\Delta \ln \frac{\bar{r}_{top}(i, o)}{\bar{r}(i, o)}$	$\Delta \text{Average Appeal}$ (reg. contr.)	$\Delta \text{Average Appeal}$ (RW)
a) Decomposition by Country-Industry Pair						
-0.03	-0.27	-0.15	0.07	0.31	0.77	0.87
b) Decomposition by Industry						
-0.30	-0.75	-0.43	0.10	0.78	0.68	0.84
c) Decomposition by Industry (only top-100 Firms)						
-0.22	0.00	-0.53	0.22	0.09	1.64	1.91

Notes: The table reports the decomposition of the share of sales by the top-4 firms in each country-industry pair (panel a), in each industry (panel b), and in each industry considering only the top-100 firms (panel c). Reported figures are simple averages computed across all country-industry pairs (panel a) and across all industries (panels b and c).

suggested by Redding and Weinstein (2017).¹⁰ For the median industry, the estimated σ_i equals 3.3 for *RW* and 4.2 for *reg. contr.*. For each value of σ_i , we compute $\tilde{\gamma}(\omega)$ separately for 2002 and 2012, following either (7) or (10) depending on the decomposition.

Table 2 shows the decomposition of the share of sales by the top-4 firms in each country-industry pair (panel a) and in each industry (panel b).¹¹ Column (1) reports the variable to be explained, i.e., the change in the log shares of top firms. Columns (2)-(5) display the contribution of each term on the right-hand side of the decompositions in (5) and (9). The coefficients are simple averages across all country-industry pairs (panel a) or across all industries (panel b). By construction, the contributions of the four components add up to the total change to be explained in the first column. Finally, columns (6) and (7) decompose the relative revenue per product of top firms (i.e., the figures reported in column 5) according to (8) and (11). Each reported coefficient measures the percentage contribution of the change in average appeal of top firm-products, i.e., the contribution of the first term in (8) and (11), for a different estimate of σ_i . Each contribution is computed by dividing the arithmetic mean of the change in average appeal across all country-industry pairs (panel a) or across all industries (panel b) by the corresponding arithmetic mean of the change in relative revenue. The contribution of heterogeneity, i.e., differential growth of appeal across firm-products, is one minus the previous coefficient and is omitted to save space.

Consider first the decomposition of $\Delta \ln s_{top}(i, o)$ and $\Delta \ln s_{top}(i)$. Table 2 shows that the by far most important factor in explaining the fall in concentration in the US import market is the extensive margin. First, there is a large increase in the number of firms that start

¹⁰We have also found similar results (available upon request) using the estimates of σ_i obtained by Broda and Weinstein (2006), who use a different estimation approach and aggregate product-level US import data for earlier years.

¹¹In the Appendix, we perform the decompositions using the share of sales by the top-3 or top-5 firms. The results are similar.

exporting to the US in a given industry ($-\Delta \ln n^f < 0$). Second, the extensive margin plays an important role also within firms. While all firms are shedding products ($-\Delta \ln n^p > 0$), top firms are dropping proportionally more products than other firms ($\Delta \ln n_{top}^p < \Delta \ln n^p$). Other things equal, the increase in the number of firms and the decrease in the relative number of products by top firms would have commanded a pervasive fall in industry concentration. On the other hand, the intensive margin has worked in the opposite direction. The average sales per product of top firms grew significantly relative to the rest ($\Delta \ln \bar{r}_{top}/\bar{r} > 0$), pushing towards rising concentration. Interestingly, all these effects are stronger when focusing on concentration from all origins: entry is stronger, but so is divergence of top firm-products (panel b). However, when considering firms from a single origin, the opposite effects of the intensive and extensive margins almost exactly cancel out (panel a).

Consider next the decomposition of the intensive margin. Table 2 shows the fraction explained by changes in the average appeal of top firm-products under alternative estimates of σ_i . This contribution ranges from 77% to 87% when comparing firms from a single origin (panel a) and from 68% to 84% when comparing firms from all origins. It implies that firms do not grow uniformly and that differential growth within top-firm products contributes significantly to raising industry concentration.

These results indicate that one reason for the fall of concentration in international markets is the sheer increase in the total number of exporting firms. But they also suggest that these new exporters are likely to be small. To neutralize this effect, we now decompose the share of sales by the top-4 firms over the top-100 firms in each industry. In this way, the extensive margin across firms is eliminated: $\Delta \ln n^f$ is zero by construction. The results are reported in panel c). Interestingly, industry concentration is still falling significantly even among the top-100 firms, due to top firms losing products relative to their competitors. Sales per product of top firms still grew relative to the rest, but by a much smaller margin. When decomposing the latter change, we now find that changes in average appeal of top firm-products explain more than 100%. In other words, top-4 firm-products are actually converging when compared to top-100 firms in the industry. This evidence suggests that while top firms are diverging from their national competitors, they are increasingly more similar to each other at the global level.

Table 3 reports robustness checks on the baseline decompositions shown in the first two panels of Table 2. In panels a) and b) we exclude countries for which the coverage of Piers is less extensive, i.e., below 45% of total exports to the US.¹² The contributions are largely

¹²These are the countries belonging to the first group of countries in Figure 2b), reported in the Appendix.

Table 3: Decomposition of the Share of Sales by the Top-4 Firms (Robustness Checks)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta \ln s_{top}(i, o)$	$-\Delta \ln n^f(i, o)$	$\Delta \ln n_{top}^p(i, o)$	$-\Delta \ln n^p(i, o)$	$\Delta \ln \frac{\bar{r}_{top}(i, o)}{\bar{r}(i, o)}$	$\Delta \text{Average Appeal}$ (reg. contr.)	$\Delta \text{Average Appeal}$ (RW)
a) Excluding Small Countries (Decomposition by Country-Industry Pair)						
-0.04	-0.29	-0.15	0.08	0.33	0.78	0.87
b) Excluding Small Countries (Decomposition by Industry)						
-0.30	-0.77	-0.43	0.11	0.79	0.72	0.85
c) Excluding Countries with Small Market Shares (Decomposition by Country-Industry Pair)						
-0.03	-0.27	-0.15	0.07	0.32	0.77	0.87
d) Excluding Countries with Small Market Shares (Decomposition by Industry)						
-0.30	-0.76	-0.43	0.11	0.79	0.68	0.84
e) Excluding Countries with Large Market Shares (Decomposition by Country-Industry Pair)						
-0.02	-0.23	-0.12	0.06	0.27	0.81	0.91
f) Excluding Countries with Large Market Shares (Decomposition by Industry)						
-0.18	-0.46	-0.43	0.10	0.61	0.91	0.96
g) Excluding Industries with High Shares of Imported Inputs (Decomposition by Country-Industry Pair)						
-0.03	-0.22	-0.15	0.07	0.27	0.78	0.89
h) Excluding Industries with High Shares of Imported Inputs (Decomposition by Industry)						
-0.30	-0.70	-0.42	0.10	0.72	0.71	0.88

Notes: In panels a) and b), small countries are those for which the share of Piers exports in total exports to the US is smaller than 45% (i.e., the first group of countries in map b of Figure 2). In panels c) and d), small market shares are those falling below the 5th percentile of the distribution of market shares in each industry and year. In panels e) and f), large market shares are those falling above the 95th percentile of the distribution of market shares in each industry and year. In panels g) and h), industries with high shares of imported inputs are those for which the average share of imports of parts and components in total US imports over 1972-2001 is above 25%. Reported figures are simple averages computed across all country-industry pairs (panels a, c, e, and g) or across all industries (panels b, d, f and h).

unchanged, suggesting that our decompositions are not driven by countries for which the coverage of Piers is more moderate. In panels c) and d), we exclude small countries whose market shares fall below the 5th percentile of the distribution of market shares in each industry and year. By contrast, in panels e) and f) we exclude large countries whose market shares fall above the 95th percentile of the same distribution. These exercises show that our decompositions are not crucially driven by either small or large countries. Finally, in panels g) and h) we exclude industries for which the share of imports of intermediate inputs in total US imports is above 25%.¹³ The contributions are largely stable also in this case, suggesting that our decompositions are not influenced by industries with intensive participation in global value chains.

4 CONCENTRATION, FIRMS AND SALES

In this section we investigate, by means of regression analysis, the correlates of the change in concentration, and its level, both by country-industry and by industry only. In particular, we focus on two concentration measures, the Herfindahl index and the share of sales by the

¹³We identify these industries using data on imports of parts and components over 1972-2001 from Schott (2004).

top-4 firms, and regress their changes on the changes in a number of variables including, most notably, the number of firms and average sales per firm.¹⁴ This exercise complements the exact decompositions in various ways. It allows us to consider alternative measures of concentration, to include various controls, and to compare sources of variation over time and in the cross-section. Rather than being aimed at identifying causal effects, it is meant to give us more confidence on our previous findings and unveil potentially interesting new stylized facts. As we will see, the regressions confirm that concentration correlates negatively with the number of firms and positively with average revenue per firm.

Table 4 reports the results for the share of top-4 firms (panel a) and the Herfindahl index (panel b) at the country-industry level. In columns (1) and (2), we regress the change in the concentration measure between 2002 and 2012 on the change in average sales per firm and in the number of firms, respectively. All variables are expressed in logarithms. Both coefficients are very precisely estimated and exhibit the expected positive and negative sign, implying that concentration increases when firms grow bigger and when the number of firms falls. However, the R-squared suggest that the number of firms plays a more important role than average sales per firm at driving concentration among exporters from the same origin country. In column (3), we use as a regressor the change in average prices per firm. While this variable has a positive and precisely estimated coefficient, its explanatory power in terms of the R-squared is limited. This result is also expected. On the one hand, top firms sell high-appeal products, which tend to be more expensive. On the other hand, it is known that prices explain little of the observed variation in sales (e.g., Hottman, Redding and Weinstein, 2016, and Bonfiglioli, Crinò and Gancia, 2018b).

In column (4), we include the three variables in the same specification. The coefficients on average sales and number of firms are largely unchanged, whereas the coefficient on average prices switches sign although it remains small. The negative sign for prices is consistent with the positive correlation between appeal and prices: holding constant sales and the number of firms, an increase in the appeal of top firms (i.e., higher concentration) must be compensated by a decrease in average appeal. In column (5), we also control for the level of concentration in 2002, and in column (6) we add country and industry fixed effects to account for heterogeneous trends in market shares across sectors (reflecting, e.g., technical change) or countries (reflecting, e.g., economic growth). The coefficient for lagged concentration is negative and significant, suggesting that concentration has been converging across country-

¹⁴See the Appendix for analogous regressions using the share of sales by the top-3 or top-5 firms as alternative measures of concentration.

Table 4: Correlates of Concentration across Country-Industry Pairs

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a) Log Share of Sales by Top-4 Firms							
	Differences	Differences	Differences	Differences	Differences	Differences	Levels
Δ ln Average Sales	0.059*** [0.002]			0.073*** [0.002]	0.072*** [0.002]	0.073*** [0.002]	
Δ ln No. of Firms		-0.214*** [0.005]		-0.225*** [0.005]	-0.228*** [0.005]	-0.192*** [0.006]	
Δ ln Average Price			0.024*** [0.005]	-0.043*** [0.004]	-0.035*** [0.004]	-0.048*** [0.004]	
Lagged ln Top Share					-0.186*** [0.014]	-0.358*** [0.015]	
ln Average Sales							0.054*** [0.001]
ln No. of Firms							-0.245*** [0.002]
ln Average Price							-0.012*** [0.001]
Country & Industry FE	No	No	No	No	No	Yes	No
Obs.	7044	7044	7044	7044	7044	7044	17528
R2	0.08	0.25	0.00	0.35	0.39	0.51	0.74
b) Log Herfindahl Index							
	Differences	Differences	Differences	Differences	Differences	Differences	Levels
Δ ln Average Sales	0.128*** [0.005]			0.179*** [0.004]	0.170*** [0.004]	0.165*** [0.004]	
Δ ln No. of Firms		-0.485*** [0.008]		-0.532*** [0.008]	-0.521*** [0.008]	-0.455*** [0.009]	
Δ ln Average Price			0.041*** [0.010]	-0.110*** [0.008]	-0.087*** [0.008]	-0.109*** [0.008]	
Lagged ln Herfindahl					-0.278*** [0.008]	-0.441*** [0.010]	
ln Average Sales							0.127*** [0.002]
ln No. of Firms							-0.536*** [0.003]
ln Average Price							-0.026*** [0.002]
Country & Industry FE	No	No	No	No	No	Yes	No
Obs.	9864	9864	9864	9864	9864	9864	24693
R2	0.08	0.26	0.00	0.39	0.47	0.56	0.74

Notes. The dependent variable is the log share of sales by the top-4 firms in panel a) and the log Herfindahl index in panel b). In the regressions in columns (1)-(6), all variables are expressed in differences between 2002 and 2012; the regression in column (7) is estimated with all variables in levels on pooled data by country and industry for the years 2002 and 2012. The standard errors, reported in square brackets, are robust to heteroskedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

Table 5: Correlates of Concentration across Industries

	(1)	(2)	(3)	(4)	(5)	(6)
Log Share of Sales by Top-4 Firms						
	Differences	Differences	Differences	Differences	Differences	Levels
$\Delta \ln$ Average Sales	0.300*** [0.028]			0.334*** [0.032]	0.326*** [0.032]	
$\Delta \ln$ No. of Firms		-0.243*** [0.047]		-0.212*** [0.050]	-0.218*** [0.052]	
$\Delta \ln$ Average Price			0.188*** [0.072]	-0.242*** [0.075]	-0.218*** [0.073]	
Lagged \ln Top Share					-0.187*** [0.058]	
\ln Average Sales						0.194*** [0.014]
\ln No. of Firms						-0.281*** [0.010]
\ln Average Price						-0.073*** [0.016]
Obs.	361	361	361	361	361	725
R2	0.23	0.05	0.02	0.27	0.29	0.61
Log Herfindahl Index						
	Differences	Differences	Differences	Differences	Differences	Levels
$\Delta \ln$ Average Sales	0.602*** [0.054]			0.667*** [0.065]	0.638*** [0.064]	
$\Delta \ln$ No. of Firms		-0.508*** [0.084]		-0.438*** [0.088]	-0.439*** [0.090]	
$\Delta \ln$ Average Price			0.400*** [0.145]	-0.477*** [0.154]	-0.400*** [0.148]	
Lagged \ln Herfindahl					-0.318*** [0.050]	
\ln Average Sales						0.364*** [0.028]
\ln No. of Firms						-0.529*** [0.018]
\ln Average Price						-0.140*** [0.032]
Obs.	366	366	366	366	366	732
R2	0.23	0.06	0.02	0.29	0.35	0.59

Notes. The dependent variable is the log share of sales by the top-4 firms in panel a) and the log Herfindahl index in panel b). In the regressions in columns (1)-(5), all variables are expressed in differences between 2002 and 2012; the regression in column (6) is estimated with all variables in levels on pooled data by 4-digit industry for the years 2002 and 2012. The standard errors, reported in square brackets, are robust to heteroskedasticity. ***, **, *: denote significance at the 1%, 5%, and 10% level, respectively.

industries, but the estimates for the other variables are unaffected. In column (7), we estimate the same specification as in column (4) in levels on pooled data. Interestingly, the same patterns found for changes also seem to hold in the cross-section. Even the point estimates are very similar in magnitude. These results are reassuring, dispelling the concern that, with only two time observations, changes could be contaminated by noise.

In Table 5, we report the estimates of the same specifications as in Table 4 at the industry level. The coefficients are significant, exhibit the same signs as in the country-industry regressions, and are also similar in magnitude. The R-squared from the univariate specifications, however, follow the opposite pattern, suggesting that changes in the number of firms matter much less than average revenue per firm in explaining concentration in the US import market. This may be due to the fact that firms entering the US import market between 2002 and 2012 tend to be small compared to established exporters and hence they have little impact on overall concentration, while reducing significantly average sales per firm. At the country level, instead, entering firms may have more impact on concentration, given that the average number of exporting firms per country-industry pair is more limited.

5 CONCLUSIONS

Much ink has been spilled on the recent increase of industrial concentration, raising concerns that the advent of giant companies may usher in an era of monopolies, growing profit shares and low economic dynamism. However, all existing evidence has been based on national data, which are not necessarily informative of the level of concentration in markets that are increasingly global. In this paper, we have documented for the first time what happened to concentration in the largest international market in the world, namely, the market of US imports. This has allowed us not only to complement national studies, but also to draw a comprehensive picture of how global firms from all countries compete in a single destination.

Our findings challenge the view that markets are becoming less competitive. Concentration of US imports by country of origin has remained stable while it has fallen significantly when pooling firms from all origins. To shed more light on this phenomenon, we have implemented a simple structural decomposition. One of the main factors behind falling concentration is the large increase in the number of firms and products exported to the United States. Pushing in the opposite direction, sales per product of top firms have increased relative to the average. Within firms, all exporters are dropping products, but top firms are doing it at a faster rate. We also find evidence of national divergence versus global convergence among top firms.

These results seem consistent with the “superstar firm hypothesis” whereby increased concentration may be the result of markets being more competitive (Van Reenen, 2018). They are also remarkably consistent with the reallocations predicted by leading models of international trade with heterogeneous multi-product firms. Hence, they suggest a possibly more benign view of concentration. Even if our data covers only imports and are limited to the manufacturing sector, they are nevertheless consistent with the complementary findings in Rossi-Hansberg, Sarte and Trachter (2018) that competition has been rising in local US markets.

However, our results also show that firms are growing more and more unequal, a finding that resonates with recent evidence using very different data.¹⁵ Some possible explanations for this widespread trend may include changes in innovation strategies (e.g., Bonfiglioli, Crinò and Gancia 2018a and 2019, Benhabib, Perla and Tonetti, 2017, König, Lorenz and Zilibotti, 2016, Dhingra, 2013) or stronger sorting between firms, suppliers and workers (e.g., Bonfiglioli and Gancia, 2019, Song, Price, Guvenen, Bloom and von Wachter, 2015). Better understanding the causes and consequences of this process is therefore an important question for future research.

¹⁵See for instance Bonfiglioli, Crinò and Gancia (2018a and 2019), Dunne, Foster, Haltiwanger and Troske (2004), Faggio, Salvanes and Van Reenen (2010).

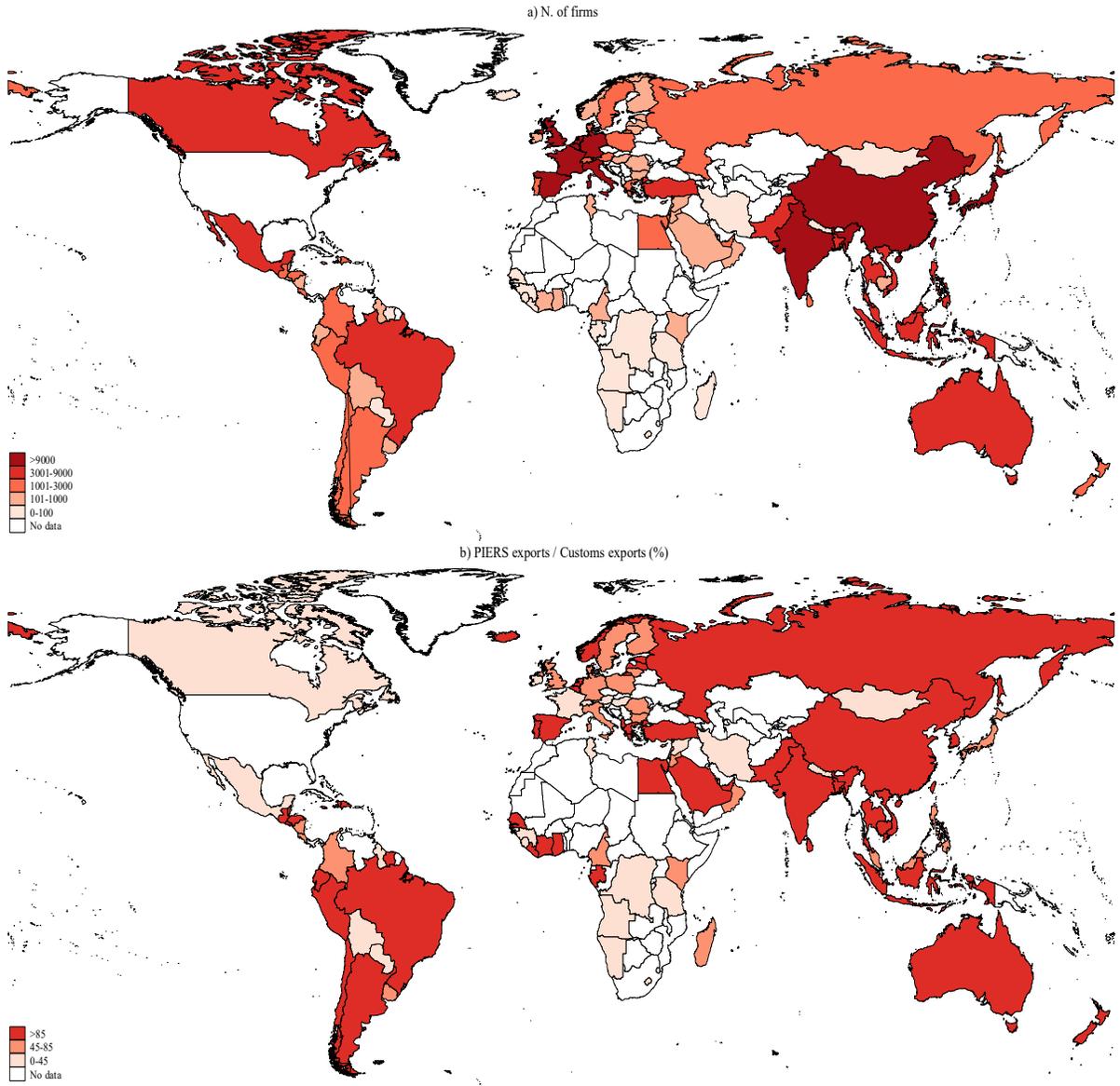
APPENDIX A THE DATA

Our analysis uses information on the universe of waterborne import transactions of the US in the years 2002 and 2012. This information comes from the Piers database administered by IHS Markit. In particular, our data set contains the complete name of each firm exporting to the US by sea in 2002 and/or 2012, its country of origin, and the total value (in US dollars) and quantity (in kilograms) of its exports by product (according to the 6-digit level of the HS classification).

Figure 2 and Table 6 provide information on the coverage of Piers. Figure 2a shows that the number of firms exporting to the US is particularly high in Canada and Mexico, Latin American countries, Europe and South-East Asia (especially China).¹⁶ Figure 2b describes the coverage of Piers in terms of export value rather than number of firms. Darker colors indicate a better coverage, as measured by the ratio between the value of total exports computed from Piers and the same value computed from customs data (Feenstra, Romalis and Schott, 2002). Although Piers registers waterborne transactions only, its coverage exceeds 80% of total exports to the US for the average country. Not surprisingly, coverage is less extensive for Canada and Mexico, which do not rely on maritime trade as their main mode of export to the United States. Nevertheless, these countries have a large number of firms exporting to the US, as shown in Figure 2a.

Table 6 provides further details on sample coverage and composition. Panel a) confirms the high coverage of Piers, showing that Piers accounts for 83% of the average country's total exports to the US and for 77% of the total exports by the median country. Feenstra and Weinstein (2017) report similar figures using an earlier and more limited version of the Piers database. Panel b) provides details on sample composition. All variables in this panel are computed separately for each country-industry-year triplet, and the reported statistics are calculated across all triplets in the data set. The average triplet has 44 firms and 55 firm-product pairs, a value of total exports to the US exceeding \$60 million and average exports

¹⁶The number of firms in Piers compares well with the same figure from other sources. For instance, across the 34 countries covered by both Piers and the World Bank Exporter Dynamics Database (EDD), which is based on information for the universe of export transactions obtained from each country's government custom agency, the coverage rate of our sample is 63% of the number of exporting firms registered in the EDD for the average or median country. Conversely, the US Customs and Border Protection database substantially overshoots the number of foreign firms exporting to the US for most countries, with an overcounting rate of 25% on average (Kamal, Krizan and Monarch, 2015). While some of the firms in Piers could be trading companies, we do not find any such company among the top-10 exporting firms in any 2-digit industry, suggesting that the majority of firms in our sample are actual exporters.



Source: Piers (IHS Markit), US import data for 2002 and 2012. Darker colors indicate a higher number of manufacturing firms exporting to the US (map a) or a higher ratio between the value of total manufacturing exports to the US obtained from Piers and the value obtained from customs data (map b). All figures are averages between 2002 and 2012.

Figure 2: Data Coverage

Table 6: Descriptive Statistics on Sample Coverage and Composition

	Mean	Median	Std. Dev.
a) <u>Sample coverage</u>			
Share of PIERS exports in total exports to the US (based on customs data)	0.83	0.77	0.55
b) <u>Sample composition</u>			
N. of firms	44	8	249
N. of firm-product pairs	55	9	316
Total exports (\$1000)	60347	2360	536000
Average exports per firm-product (\$1000)	1273	230	11058

Notes. The variable in panel a) is computed for each country in the years 2002 and 2012. Reported statistics are the mean, median and standard deviation of this variable across all countries and years. The variables in panel b) are computed for each country-industry-year triplet. Reported statistics are the mean, median and standard deviation of these variables across all triplets.

per firm-product slightly above \$1 million.

The fact that all firms in Piers use the same export mode (by sea) favors comparability. Compared to customs data from the US Customs and Border Protection, the Piers data are slightly less detailed in terms of product classification (6-digit vs. 10-digit) but contain the full name of each firm, thereby allowing us to precisely identify firms and reduce the risk of over-counting them. To match and aggregate firms appearing in the data more than once with similar names, we use a string matching algorithm that first homogenizes standard expressions (e.g., it converts the extensions "Lim." and "LTD" in "Limited") and then exploits the Levenshtein edit distance to match firms.¹⁷ Then, we assign products to industries by mapping each HS6 product code exported by a firm to a 4-digit SIC industry, using a correspondence table developed by the World Integrated Trade Solutions.

We further clean the data to mitigate the risk of including transactions contaminated by reporting mistakes. In particular, we drop firms with obvious inconsistencies in their name. We further drop observations corresponding to firms that, in a given industry and year, have total exports to the US below \$1,000. We also exclude observations corresponding to firms that, in a given industry and year, have unit values for their products above the top or below the bottom 0.01% of the unit value distribution for that year. Finally, we exclude country-industry-year triplets with less than two varieties exported to the US, since we cannot compute dispersion measures for these triplets. Our final data set comprises 1,311,835 observations at the firm-product-year level. Firms belong to 366 manufacturing industries and 104 origin countries spanning the five continents.

¹⁷The algorithm computes the Levenshtein edit distance between all pairwise combinations of firm names sharing the same first character. The distance is then normalized by the length of the longest string and a match is formed if the normalized edit distance is below a 5% threshold.

APPENDIX B ESTIMATING THE ELASTICITY OF SUBSTITUTION

To perform our decomposition exercises, we need estimates of the elasticity of substitution between varieties. Using our data, we estimate elasticities following two alternative methodologies, based on a set of assumptions on the demand side. Our main assumption is that preferences over varieties ω in industry i take the constant elasticity of substitution form as in (1), which implies that sales of variety ω in sector i can be expressed as in (3).

B.1 THE REVERSE-WEIGHTING ESTIMATOR

Following Redding and Weinstein (2017), we focus on the price index (2) for each industry i , and start by obtaining three equivalent expressions for the change in this price index between 2002 and 2012. Dropping industry labels to save on notation, and using the time subscript t when referring to the year 2012 and $t - 1$ when referring to 2002, these expressions read as follows:

$$\frac{P_t}{P_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right)^{\frac{1}{\sigma-1}} \Theta_{t-1,t}^F \left\{ \sum_{\omega \in \Omega_{t,t-1}} s_{t-1}^*(\omega) \left[\frac{p_t(\omega)}{p_{t-1}(\omega)} \right]^{1-\sigma} \right\}^{\frac{1}{1-\sigma}}, \quad (12)$$

$$\frac{P_t}{P_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right)^{\frac{1}{\sigma-1}} (\Theta_{t,t-1}^B)^{-1} \left\{ \sum_{\omega \in \Omega_{t,t-1}} s_t^*(\omega) \left[\frac{p_t(\omega)}{p_{t-1}(\omega)} \right]^{-(1-\sigma)} \right\}^{-\frac{1}{1-\sigma}}, \quad (13)$$

$$\frac{P_t}{P_{t-1}} = \left(\frac{\lambda_{t,t-1}}{\lambda_{t-1,t}} \right)^{\frac{1}{\sigma-1}} \frac{\tilde{P}_t^*}{\tilde{P}_{t-1}^*} \left(\frac{\tilde{S}_t^*}{\tilde{S}_{t-1}^*} \right)^{\frac{1}{\sigma-1}}, \quad (14)$$

where $\Omega_{t,t-1}$ denotes the set of common varieties in both years; $s^*(\omega)$ denotes the share of common variety ω in expenditure on all common varieties; \tilde{S}^* and \tilde{P}^* denote the geometric averages of $s^*(\omega)$ and $p(\omega)$, respectively, computed on common varieties; $(\lambda_{t,t-1}/\lambda_{t-1,t})^{1/(\sigma-1)}$ is the variety-adjustment term, which adjusts the common varieties price index for entering

and exiting varieties; and

$$\Theta_{t-1,t}^F \equiv \left\{ \frac{\sum_{\omega \in \Omega_{t,t-1}} s_{t-1}^*(\omega) \left[\frac{p_t(\omega)}{p_{t-1}(\omega)} \right]^{1-\sigma} \left[\frac{\gamma_t(\omega)}{\gamma_{t-1}(\omega)} \right]^{\sigma-1}}{\sum_{\omega \in \Omega_{t,t-1}} s_{t-1}^*(\omega) \left[\frac{p_t(\omega)}{p_{t-1}(\omega)} \right]^{1-\sigma}} \right\}^{\frac{1}{1-\sigma}} = \left\{ \sum_{\omega \in \Omega_{t,t-1}} s_t^*(\omega) \left[\frac{\gamma_{t-1}(\omega)}{\gamma_t(\omega)} \right]^{\sigma-1} \right\}^{\frac{1}{\sigma-1}}$$

$$\Theta_{t,t-1}^B \equiv \left\{ \frac{\sum_{\omega \in \Omega_{t,t-1}} s_t^*(\omega) \left[\frac{p_{t-1}(\omega)}{p_t(\omega)} \right]^{1-\sigma} \left[\frac{\gamma_{t-1}(\omega)}{\gamma_t(\omega)} \right]^{\sigma-1}}{\sum_{\omega \in \Omega_{t,t-1}} s_t^*(\omega) \left[\frac{p_{t-1}(\omega)}{p_t(\omega)} \right]^{1-\sigma}} \right\}^{\frac{1}{1-\sigma}} = \left\{ \sum_{\omega \in \Omega_{t,t-1}} s_{t-1}^*(\omega) \left[\frac{\gamma_t(\omega)}{\gamma_{t-1}(\omega)} \right]^{\sigma-1} \right\}^{\frac{1}{\sigma-1}}$$

are the forward and backward differences of the price index, which evaluate its change using varieties' expenditure shares in $t - 1$ and t , respectively.

The three ways of expressing the change in the price index are equivalent. However, the formulation in (14) is the only one that exclusively depends on prices and expenditure shares, and not also on the demand parameters γ (i.e., this formulation is money-metric). Note also that the three expressions depend on the elasticity of substitution, σ . Hence, the idea of the RW estimator is to search the value of σ that renders the three expressions for the change in the price index consistent with the same money-metric utility function. This requires imposing the following identifying assumption:

$$\Theta_{t-1,t}^F = (\Theta_{t,t-1}^B)^{-1} = 1, \quad (15)$$

which means that changes in γ over time average out.

Combining (12)-(14) and using (15), one can construct a generalized method of moment estimator for σ . In particular, the following moment functions obtain:

$$M(\sigma) \equiv \begin{pmatrix} m_t^F(\sigma) \\ m_t^B(\sigma) \end{pmatrix} \equiv \begin{pmatrix} \frac{1}{1-\sigma} \ln \left\{ \sum_{\omega \in \Omega_{t,t-1}} s_{t-1}^*(\omega) \left[\frac{p_t(\omega)}{p_{t-1}(\omega)} \right]^{1-\sigma} \right\} - \ln \left[\frac{\tilde{P}_t^*}{\tilde{P}_{t-1}^*} \left(\frac{\tilde{S}_t^*}{\tilde{S}_{t-1}^*} \right)^{\frac{1}{\sigma-1}} \right]}{-\frac{1}{1-\sigma} \ln \left\{ \sum_{\omega \in \Omega_{t,t-1}} s_t^*(\omega) \left[\frac{p_t(\omega)}{p_{t-1}(\omega)} \right]^{-(1-\sigma)} \right\} - \ln \left[\frac{\tilde{P}_t^*}{\tilde{P}_{t-1}^*} \left(\frac{\tilde{S}_t^*}{\tilde{S}_{t-1}^*} \right)^{\frac{1}{\sigma-1}} \right]} \end{pmatrix} = \begin{pmatrix} -\ln \Theta_{t-1,t}^F \\ \ln \Theta_{t,t-1}^B \end{pmatrix}.$$

The RW estimator $\hat{\sigma}^{RW}$ solves:

$$\hat{\sigma}^{RW} = \arg \min \left\{ M(\hat{\sigma}^{RW})' \times \mathbb{I} \times M(\hat{\sigma}^{RW}) \right\},$$

where \mathbb{I} is the identity matrix. Because the two moments are weighted by the identity matrix, the RW estimator minimizes the sum of squared deviations of the aggregate demand parameters $((-\ln \Theta_{t-1,t}^F)^2 + (\ln \Theta_{t,t-1}^B)^2)$ from zero. Hence, the RW estimator selects the value of σ that minimizes the squared deviations of the forward and backward differences of the price index from a money-metric utility function.

B.2 AN ALTERNATIVE ESTIMATION

As an alternative approach, we identify the elasticity of substitution out of the variation in the variance of sales across varieties ω sold from different origin countries o in each sector i . Starting from the expression for sales in equation (3), we obtain the variance of log sales across varieties sold from country o in industry i as

$$\mathbb{V}(\ln r(i, o)) = (\sigma_i - 1)^2 \mathbb{V}(\ln \tilde{\gamma}(i, o)).$$

Taking logs and adding time subscripts yields

$$\ln \mathbb{V}(\ln r_t(i, o)) = \overbrace{2 \ln(\sigma_i - 1)}^{\text{Industry fixed effect, } \alpha_i} + \ln \mathbb{V}(\ln \tilde{\gamma}_t(i, o)), \quad (16)$$

Intuitively, a higher substitutability generates more dispersion in sales for a given distribution of attributes. The limitation of this strategy is that the industry fixed effect, α_i , identifies any common component of sales dispersion across countries in a given industry, and not just the demand parameter we are interested in. On the other hand, the advantage is that purging sales dispersion of any common component across countries allows us to isolate the cross-country variation in attributes. Hence, it is a way to study heterogeneity in attributes *relative* to other countries, rather than its *absolute* level.

Since the second term in (16) cannot be observed directly, we proxy for $\ln \mathbb{V}(\ln \tilde{\gamma}_t(i, o))$ with a number of observable variables. In particular, we estimate the following specification:

$$\ln \mathbb{V}(\ln r_t(i, o)) = \alpha_i + \beta_1 \ln \mathbb{V}(\ln p_t(i, o)) + \beta_2 \ln n_t(i, o) + \nu_{ot} + \varepsilon_{oi,t}, \quad (17)$$

where $\mathbb{V}(\ln p_t(i, o))$ is the variance of log prices at time t , $\ln n_t(i, o)$ is the log number of firm-products per country-industry-year triplet, and ν_{ot} are country-time fixed effects. While prices are just one component of $\tilde{\gamma}$, controlling for their variance would be sufficient if there is a one-to-one mapping between quality and prices, as in several models of endogenous

Table 7: Descriptive Statistics on the Elasticity of Substitution

	Mean	Median	Std. Dev.	N. of Industries
Regression, controls (<i>reg. contr.</i>)	4.22	4.22	0.46	366
Reverse weighting (<i>RW</i>)	3.71	3.30	1.74	232

Notes. The statistics are computed across all industries with available information on a given elasticity of substitution. Elasticities smaller than 1 or greater than 10 are excluded.

quality. This choice is also supported by the evidence that prices are indeed a good proxy for quality (see Hottman, Redding and Weinstein, 2016, and Johnson, 2012). We choose the second proxy, $\ln n_t(i, o)$, based on previous evidence (Bonfiglioli, Crinò and Gancia, 2018a and 2019) showing that the variance of sales may vary systematically with the number of observations over which it is computed. Finally, the country-time fixed effects, ν_{ot} , make sure that the industry fixed effects, α_i , are identified from deviations of sales dispersion from its country-year means, and are not contaminated by time-varying country characteristics that could affect sales dispersion uniformly across industries, e.g., by systematically inducing some countries to specialize in high- or low-dispersion industries. After estimating (17), we retrieve σ_i as:

$$\hat{\sigma}_i^{reg.contr.} = \exp\left(\frac{\hat{\alpha}_i}{2}\right) + 1.$$

Table 7 provides descriptive statistics on the estimated σ_i . For the median industry in our sample, the estimated elasticity of substitution equals 3.3 in the case of *RW* and 4.2 in the case of *reg. contr.*. These numbers are in the same ballpark as the estimates obtained in previous studies. For instance, across the industries in our sample, the median value of the elasticity estimated by Broda and Weinstein (2006) using US import data across 10-digit products for an earlier period equals 3.1.

APPENDIX C ADDITIONAL ROBUSTNESS CHECKS

In this section, we perform additional robustness checks on our decompositions and regressions. In Table 8, we repeat the baseline decompositions reported in panels a) and b) of Table 2 using alternative measures of concentration. In particular, we use the share of sales by the top-3 or top-5 firms instead of the share of sales by the top-4 firms. The results are very similar to the baseline decompositions, suggesting that the latter are not driven by the measure of concentration.

In Tables 9 and 10, we repeat the regressions shown in Tables 4 and 5 using the log share of sales by the top-3 or top-5 firms as the dependent variable. The estimated coefficients have

Table 8: Decomposition of the Share of Sales by the Top-3 and Top-5 Firms

(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta \ln s_{top}(i, o)$	$-\Delta \ln n^f(i, o)$	$\Delta \ln n_{top}^p(i, o)$	$-\Delta \ln n^p(i, o)$	$\Delta \ln \frac{\bar{r}_{top}(i, o)}{\bar{r}(i, o)}$	$\Delta \text{Average Appeal}$ (reg. contr.)	$\Delta \text{Average Appeal}$ (RW)
a) Top-3 Firms (Decomposition by Country-Industry Pair)						
-0.03	-0.26	-0.14	0.07	0.30	0.81	0.89
b) Top-3 Firms (Decomposition by Industry)						
-0.32	-0.76	-0.45	0.10	0.80	0.79	0.90
c) Top-5 Firms (Decomposition by Country-Industry Pair)						
-0.04	-0.28	-0.15	0.08	0.32	0.74	0.85
d) Top-5 Firms (Decomposition by Industry)						
-0.28	-0.76	-0.40	0.10	0.78	0.66	0.82

Notes: Panels a) and b) report the decomposition of the share of sales by the top-3 firms in each country-industry pair and in each industry, respectively. Panels c) and d) report the same decompositions for the share of sales by the top-5 firms. Reported figures are simple averages computed across all country-industry pairs (panels a and c) or across all industries (panels b and d).

the same sign as the baseline estimates and are close to them in terms of size. This further suggests that our evidence is robust to the use of alternative measures of concentration.

Table 9: Correlates of Concentration across Country-Industry Pairs (Robustness Checks)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
a) Log Share of Sales by Top-3 Firms							
	Differences	Differences	Differences	Differences	Differences	Differences	Levels
$\Delta \ln$ Average Sales	0.061*** [0.002]			0.078*** [0.002]	0.076*** [0.002]	0.076*** [0.002]	
$\Delta \ln$ No. of Firms		-0.229*** [0.005]		-0.243*** [0.005]	-0.244*** [0.005]	-0.205*** [0.006]	
$\Delta \ln$ Average Price			0.022*** [0.005]	-0.049*** [0.004]	-0.039*** [0.004]	-0.051*** [0.004]	
Lagged \ln Top Share					-0.225*** [0.012]	-0.400*** [0.014]	
\ln Average Sales							0.058*** [0.001]
\ln No. of Firms							-0.262*** [0.002]
\ln Average Price							-0.012*** [0.001]
Country & Industry FE	No	No	No	No	No	Yes	No
Obs.	8037	8037	8037	8037	8037	8037	20016
R2	0.07	0.23	0.00	0.34	0.39	0.51	0.73
b) Log Share of Sales by Top-5 Firms							
	Differences	Differences	Differences	Differences	Differences	Differences	Levels
$\Delta \ln$ Average Sales	0.057*** [0.002]			0.070*** [0.002]	0.069*** [0.002]	0.070*** [0.002]	
$\Delta \ln$ No. of Firms		-0.203*** [0.005]		-0.212*** [0.005]	-0.216*** [0.005]	-0.187*** [0.007]	
$\Delta \ln$ Average Price			0.025*** [0.005]	-0.040*** [0.004]	-0.033*** [0.004]	-0.046*** [0.004]	
Lagged \ln Top Share					-0.152*** [0.015]	-0.314*** [0.017]	
\ln Average Sales							0.051*** [0.001]
\ln No. of Firms							-0.233*** [0.002]
\ln Average Price							-0.012*** [0.001]
Country & Industry FE	No	No	No	No	No	Yes	No
Obs.	6310	6310	6310	6310	6310	6310	15728
R2	0.08	0.25	0.00	0.36	0.39	0.51	0.74

Notes. The dependent variable is the log share of sales by the top-3 firms in panel a) and by the top-5 firms in panel b). In the regressions in columns (1)-(6), all variables are expressed in differences between 2002 and 2012; the regression in column (7) is estimated with all variables in levels on pooled data by country and industry for the years 2002 and 2012. The standard errors, reported in square brackets, are robust to heteroskedasticity. ***, **, *: denote significance at the 1%, 5%, and 10% level, respectively.

Table 10: Correlates of Concentration across Industries (Robustness Checks)

	(1)	(2)	(3)	(4)	(5)	(6)
Log Share of Sales by Top-3 Firms						
	Differences	Differences	Differences	Differences	Differences	Levels
Δ ln Average Sales	0.322*** [0.031]			0.364*** [0.036]	0.352*** [0.035]	
Δ ln No. of Firms		-0.262*** [0.049]		-0.236*** [0.052]	-0.233*** [0.054]	
Δ ln Average Price			0.209*** [0.080]	-0.276*** [0.084]	-0.235*** [0.081]	
Lagged ln Top Share					-0.253*** [0.058]	
ln Average Sales						0.204*** [0.015]
ln No. of Firms						-0.291*** [0.010]
ln Average Price						-0.079*** [0.017]
Obs.	364	364	364	364	364	730
R2	0.22	0.05	0.02	0.27	0.30	0.59
Log Share of Sales by Top-5 Firms						
	Differences	Differences	Differences	Differences	Differences	Levels
Δ ln Average Sales	0.282*** [0.025]			0.310*** [0.030]	0.306*** [0.029]	
Δ ln No. of Firms		-0.227*** [0.044]		-0.185*** [0.048]	-0.188*** [0.049]	
Δ ln Average Price			0.178*** [0.064]	-0.218*** [0.068]	-0.200*** [0.067]	
Lagged ln Top Share					-0.144** [0.058]	
ln Average Sales						0.184*** [0.013]
ln No. of Firms						-0.273*** [0.009]
ln Average Price						-0.071*** [0.015]
Obs.	360	360	360	360	360	724
R2	0.24	0.05	0.02	0.28	0.29	0.64

Notes. The dependent variable is the log share of sales by the top-3 firms in panel a) and by the top-5 firms in panel b). In the regressions in columns (1)-(5), all variables are expressed in differences between 2002 and 2012; the regression in column (6) is estimated with all variables in levels on pooled data by 4-digit industry for the years 2002 and 2012. The standard errors, reported in square brackets, are robust to heteroskedasticity. ***, **, * denote significance at the 1%, 5%, and 10% level, respectively.

REFERENCES

- [1] Anderson, Eric, Sergio Rebelo, and Arlene Wong (2018). “Markups Across Space and Time,” *NBER Working Paper No. 24434*.
- [2] Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen (2017). “The Fall of the Labor Share and the Rise of Superstar Firms,” *NBER Working Paper No. 23396*.
- [3] Baqaee, David Rezza and Emmanuel Farhi (2017). “Productivity and Misallocation in General Equilibrium,” *NBER Working Paper No. 24007*.
- [4] Barkai, Simcha (2017). “Declining Labor and Capital Shares,” *LBS Working Paper*.
- [5] Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta (2013). “Cross-Country Differences in Productivity: The Role of Allocation and Selection,” *American Economic Review*, 103, 305–334.
- [6] Benhabib, Jess, Jesse Perla and Christopher Tonetti (2017). “Reconciling Models of Diffusion and Innovation: A Theory of the Productivity Distribution and Technology Frontier,” *NBER Working Paper No. 23095*.
- [7] Bernard, Andrew B., Stephen J. Redding and Peter K. Schott (2011). “Multiproduct Firms and Trade Liberalization,” *Quarterly Journal of Economics*, 126(3), 1271–1318.
- [8] Bernard, Andrew, Brad Jensen, Stephen Redding and Peter Schott (2018). “Global Firms,” *Journal of Economic Literature*, forthcoming .
- [9] Bloom, Nicholas, Raffaella Sadun and John Van Reenen (2016). “Management as a Technology?” *NBER Working Paper No. 22327*.
- [10] Bonfiglioli, Alessandra, Rosario Crinò and Gino Gancia (2018a). “Betting on Exports: Trade and Endogenous Heterogeneity,” *Economic Journal*, 128, 612–651.
- [11] Bonfiglioli, Alessandra, Rosario Crinò and Gino Gancia (2018b). “Firms and Economic Performance: A View from Trade,” *CEPR Discussion Paper 12829*.
- [12] Bonfiglioli, Alessandra, Rosario Crinò and Gino Gancia (2019). “Trade, Finance and Endogenous Heterogeneity,” *Journal of the European Economic Association*, 17(1), 79–130.
- [13] Bonfiglioli, Alessandra and Gino Gancia (2019). “Heterogeneity, Selection and Labor Market Disparities,” *Review of Economic Dynamics*, 31, 305–325.
- [14] Broda, Christian and David Weinstein (2006). “Globalization and the Gains from Variety,” *Quarterly Journal of Economics*, 121, 541–585.
- [15] Calligaris, Sara, Chiara Criscuolo and Luca Marcolin (2018). “Mark-ups in the digital era”, OECD Science, Technology and Industry Working Papers, OECD Publishing, Paris.

- [16] De Loecker, Jan and Jan Eeckhout (2017). “The Rise of Market Power and the Macroeconomic Implications,” *NBER Working Paper No 23687*.
- [17] De Loecker, Jan and Jan Eeckhout (2018). “Global Market Power,” *NBER Working Paper No. 24768*.
- [18] Dhingra, Swati (2013). “Trading Away Wide Brands for Cheap Brands,” *American Economic Review*, 103(6): 2554-84.
- [19] Diez, Federico, Daniel Leigh, and Suchanan Tambunlertchai (2018). “Global Market Power and its Macroeconomic Implications,” *IMF Working Paper 18/137*.
- [20] di Giovanni, Julian, Andrei Levchenko and Romain Ranciere (2011). “Power Laws in Firm Size and Openness to Trade: Measurement and Implications.” *Journal of International Economics*, 85:1, 42-52.
- [21] Dunne, Timothy, Lucia Foster, John Haltiwanger and Kenneth Troske (2004). “Wage and Productivity Dispersion in US Manufacturing: The Role of Computer Investments.” *Journal of Labor Economics*, 22: 397-429.
- [22] Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu (2018). “How Costly are Markups?” *NBER Working Paper No. 24800*.
- [23] Epifani, Paolo and Gino Gancia (2011). “Trade, Markup Heterogeneity and Misallocations,” *Journal of International Economics*, 83, 1–13.
- [24] Faggio, Giulia, Kjell Salvanes and John Van Reenen (2010). “The Evolution of Inequality in Productivity and Wages: Panel Data Evidence.” *Industrial and Corporate Change*, 19: 1919-1951.
- [25] Feenstra, Robert, John Romalis and Peter Schott (2002). “U.S. Imports, Exports and Tariff Data, 1989–2001.” *NBER Working Paper No. 9387*.
- [26] Feenstra, Robert and David Weinstein (2017). “Globalization, Competition, and U.S. Welfare,” *Journal of Political Economy*, 125, 1041–1074.
- [27] Freund, Caroline and Martha Denisse Pierola (2015). “Export Superstars,” *Review of Economics and Statistics*, 97, 1023–1032.
- [28] Freund, Caroline and Dario Sidhu (2017). “Global Competition and the Rise of China,” *PIIE Working Paper 17-3*.
- [29] Fernandes, Ana, Peter Klenow, Sergii Meleshchuk, Martha Denisse Pierola and Andres Rodriguez-Clare (2018). “The Intensive Margin in Trade,” *NBER Working Paper No. 25195*.
- [30] Gaubert, Cecile and Oleg Itskhoki (2018). “Granular Comparative Advantage,” *NBER Working Paper No. 24807*.

- [31] Gennaioli, Nicola, Rafael La Porta, Florencio López-de-Silanes and Andrei Shleifer (2013). “Human Capital and Regional Development,” *Quarterly Journal of Economics*, 128(1), 105-164.
- [32] Gutierrez, German and Thomas Philippon (2018) “How EU markets became more competitive than US markets” *NBER Working Paper No. 24700*.
- [33] Hall, Robert E (2018). “New Evidence on the Markup of Prices Over Marginal Costs and the Role of Mega-Firms in the US Economy,” *NBER Working Paper No 24574*.
- [34] Hottman, Colin, Stephen Redding and David Weinstein (2016). “Quantifying the Sources of Firm Heterogeneity,” *Quarterly Journal of Economics*, 131, 1291–1364.
- [35] Hummels, David and Peter Klenow (2005). “The Variety and Quality of a Nation’s Exports,” *American Economic Review*, 95, 704–723.
- [36] Johnson, Robert (2012). “Trade and Prices with Heterogeneous Firms,” *Journal of International Economics*, 86, 43–56.
- [37] Kamal, Fariha, C.J. Krizan and Ryan Monarch (2015). “Identifying Foreign Suppliers in U.S. Merchandise Import Transactions,” *Board of Governors of the Federal Reserve System - International Finance Discussion Paper No. 1142*.
- [38] Karabarbounis, Loukas, and Brent Neiman (2018). “Accounting for Factorless Income,” *NBER Working Paper No. 24404*.
- [39] König, Michael D., Jan Lorenz and Fabrizio Zilibotti (2016). “Innovation vs. imitation and the evolution of productivity distributions,” *Theoretical Economics*, 11, 1053–1102.
- [40] Melitz, Marc (2003). “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity.” *Econometrica*, 71, 1695–1725.
- [41] Melitz, Marc, Thierry Mayer, and Gianmarco IP Ottaviano (2014). “Market Size, Competition, and the Product Mix of Exporters.” *American Economic Review* 104 (2): 495-536.
- [42] Melitz, Marc and Stephen Redding (2014). “Heterogeneous Firms and Trade,” in *Handbook of International Economics* (Gopinath, Helpman and Rogoff, eds.), Elsevier, Volume 4, 1–54.
- [43] Poschke, Markus (2015). “The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology,” *IZA Discussion Paper No. 7991*.
- [44] Redding, Stephen and David Weinstein (2017). “A Unified Approach to Aggregate Price and Welfare Measurement,” *NBER Working Paper No. 22479*.
- [45] Redding, Stephen and David Weinstein (2018). “Accounting for Trade Patterns,” *Mimeo*, Princeton University.
- [46] Rossi-Hansberg, Esteban, Pierre-Daniel Sarte and Nicholas Trachter (2018). “Diverging Trends in National and Local Concentration,” *CEPR Discussion Paper No. 13174*.

- [47] Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom and Till von Wachter (2015). “Firming Up Inequality.” *NBER Working Paper No. 21199*.
- [48] Van Reenen, John (2018). “Increasing Differences between firms: Market Power and the Macro-Economy,” *Jackson Hole conference proceedings*.