# Firm Size, Quality Bias and Import Demand*

Joaquin Blaum† Claire Lelarge‡ Michael Peters§

May 9, 2019

**Abstract**

Commonly used firm-based models of importing imply that firm productivity should have no effect on the allocation of expenditure across a common set of sourcing countries. Using French data, we show that this homotheticity property is soundly rejected: larger firms concentrate their import spending on their top varieties, holding the sourcing strategy fixed. To rationalize this finding, we propose a novel model of importing that features (i) a complementarity between firm productivity and input quality and (ii) heterogeneity across countries in their ability to produce high quality inputs. This model implies that large firms bias their spending towards countries with a comparative advantage in producing high quality inputs and hence generates a non-homothetic import demand system. We provide empirical support for this and other predictions of this theory.

*JEL Codes*: F11, F12, F14, D21, D22, D24.
*Keywords*: trade in intermediate inputs, firm heterogeneity, firm size, non-homothetic import demand, quality-complementarity.

*We thank Costas Arkolakis, Lorenzo Caliendo, Arnaud Costinot, Penny Goldberg, Oleg Itskhoki, Amit Khandelwal, Sam Kortum, Andrei Levchenko, Peter Schott and seminar participants at the University of Chicago, MIT, Yale, Brown and the NBER SI ITI meeting. We thank Pablo Fajgelbaum for his discussion at the NBER SI ITI. Previous versions of this paper circulated under the titles "Non-Homothetic Import Demand: Firm Productivity and Quality Bias" and "Firm Size and the Intensive Margin of Import Demand".

†Brown University. Corresponding author. jblaum@gmail.com. Mailing address: Department of Economics, Brown University, 64 Waterman Street, Providence, RI 02912.

‡University of Paris-Saclay, RITM, CREST and CEPR. claire.lelarge@gmail.com

§Yale University and NBER. m.peters@yale.edu

# 1 Introduction

Understanding the behavior of individual firms has been at the center of the recent research agenda in international trade. Most of this research has focused on the role of firms as exporters. Recently, there has been a new emphasis on the behavior of firms as importers of intermediate inputs, which account for two-thirds of world trade (Johnson and Noguera (2017), Miroudot et al. (2009)). This literature has for example focused on the link between imported inputs and firm productivity (Halpern et al. (2015), Blaum et al. (2018)), the effect of large crises on aggregate productivity (Gopinath and Neiman (2014), Blaum (2019)), domestic product scope (Goldberg et al. (2010)), and firms' global sourcing patterns (Antràs et al. (2017)).

While different in focus, the contributions in the literature have mostly relied on a common theoretical framework. Because productivity is assumed to be factor neutral, this common framework features a strong restriction on firms' import demand system: holding the set of sourcing countries fixed, firm productivity should not affect the allocation of spending. In other words, essentially all existing firm-based models of input trade imply that firms' import demand is homothetic.

In this paper, we use French micro data to test this homotheticity prediction and show that it is systematically rejected. We find that firms differ substantially in how they allocate their spending across a common set of sourcing countries and that these differences are systematically related to firm size. In particular, large firms concentrate their expenditure on their most important trading partners. To rationalize this finding, we extend the baseline framework in the literature by incorporating (i) a complementarity between input quality and firm productivity and (ii) heterogeneity across countries in their ability to produce high quality inputs. This extended model predicts that import demand is non-homothetic because large firms import expensive, high quality inputs and therefore bias their spending patterns towards countries with a comparative advantage in the production of inputs of higher quality. We rely on data on unit values to verify this and other predictions of the extended model.

We start our analysis by laying out a framework that nests most of the contributions in the literature on input trade, which we refer to as the baseline model of importing. Firms can source their inputs from multiple foreign countries, there is imperfect substitutability, and the production technology features constant returns to scale given the set of inputs sourced. Importantly, as long as firm productivity is factor neutral, this baseline framework features a homothetic import demand system. In particular, holding the set of sourcing countries fixed, firms' expenditure shares depend only on the prices and qualities of the sourcing countries and should therefore be independent of firm size.

We next implement a test for this prediction in the data. As stressed by the theory, this requires appropriately controlling for the extensive margin of trade. We do so by comparing firms sourcing a particular product from the exact same set of countries.[1] We find systematic violations of the baseline model as firms disagree substantially in their allocation of import expenditure across countries. Importantly, such disagreement is not merely idiosyncratic but is systematically related to firm size. In particular, we show that the expenditure distribution of large firms first-order stochastically dominates that of smaller firms. In addition, the allocation of domestic vs foreign expenditure is also non-homothetic. Large firms spend relatively more on domestic inputs, after controlling for the extensive margin.

Our findings are at odds with a large class of models. The reason is that the homotheticity of import

[1] In contrast, a common approach in the literature has been to rely on the number of sourcing countries to control for the extensive margin of trade. This is accurate only under special parameterizations of the model. We show empirically that the number is not a sufficient statistic for the set of sourcing countries.

demand in the baseline model is independent of many aspects of the theory. Most importantly, we do not have to impose any restriction on firms' extensive margin of importing. Hence, our test applies regardless of whether importing is subject to fixed costs, to what extent such fixed costs vary across firms and countries, or whether firms select their trading partners through a dynamic process of network formation. Moreover, we also allow firms to compete in output markets in an unrestricted way and we do not need to impose particular functional forms for firms' production technologies.

We then propose a tractable extension of the baseline model of importing that endogenously generates a non-homothetic import demand and can rationalize our findings. We do so by incorporating two additional ingredients. First, every sourcing country offers a menu of inputs which are differentiated by quality. Higher quality inputs are more costly to produce and, crucially, countries differ in the elasticity of production costs with respect to quality. This country-specific elasticity controls the degree of comparative advantage in the production of high quality inputs. Second, we depart from firm productivity being factor neutral and allow for a complementarity between firm productivity and input quality as in Kugler and Verhoogen (2012).

We derive two results in the extended model. First, more productive firms buy higher quality inputs within any given country. This follows from the complementarity between input quality and firm productivity. Second, large firms particularly benefit from countries that produce quality more efficiently, hence generating an import demand system that is non-homothetic at the country level. This non-homotheticity arises from the interaction between the quality-productivity complementarity at the firm-level and the heterogeneity in quality production costs across countries. Importantly, without differences in comparative advantage across countries, the within-country sorting pattern of firm productivity and input quality would end up being irrelevant for the allocation of expenditure across countries, generating an import demand system that is homothetic at the country level as in the baseline model.

We then provide empirical support for the extended model using firm-level data on input prices. We proceed in three steps. First, we document that larger firms pay higher prices for their inputs within origin countries. Second, we use this within-country relationship between input prices and firm size to estimate the elasticity of production costs with respect to quality at the country level. Intuitively, because the theory implies that large firms buy higher quality inputs, we can infer unobserved input quality from the size of its buyer. Thus, we can recover the price-quality elasticity by comparing the prices paid by firms of different size within sourcing countries. Importantly, we find substantial cross-country heterogeneity in the estimated elasticities and hence in comparative advantage in quality production.[2]

Finally, in a third step, we test for the main prediction of the extended model that large firms bias their spending towards countries that efficiently produce high quality inputs. Using the price-quality elasticities estimated in the previous step, we find empirical support for this prediction. In addition, we show that countries whose price-quality elasticity was estimated to be low tend to account for large shares of firms' import spending. Taken together, these results show that the model of input quality choice generates a non-homothetic import demand system that is in line with the data.

**Related Literature.** This paper is most closely related to the recent literature on firm-based models of importing. This literature takes a structural approach to study firms' importing decisions and explores a variety of questions. Examples include the relationship between firms' import activities and firm and aggregate productivity (see. e.g. Halpern et al. (2015), Blaum et al. (2018), Goldberg et al. (2010), Ramanarayanan

[2]We also show that these differences in comparative advantage are positively correlated with income per capita and hence systematically related to aggregate productivity at the country level.

(2014), Gopinath and Neiman (2014), Kasahara and Rodrigue (2008)), the analysis of firms' sourcing decisions (Antràs et al. (2017), Lu et al. (2016)), or the exchange-rate disconnect (Amiti et al. (2014)). Despite their different focus, all of these papers employ frameworks where the resulting import demand system is homothetic at the intensive margin.[3] We present systematic evidence against this implication and therefore reject a property which is shared by many contributions in the literature.

To explain the observed non-homotheticity of firms' import demand, we build a model of importing where firms can choose the quality of their inputs within sourcing countries. To the best of our knowledge, we are the first to provide a theory of importing with such firm-level non-homotheticities. We follow Kugler and Verhoogen (2012) in assuming that input quality and firm productivity are complements. In contrast to them, we focus on the choice of input quality within foreign countries and how this affects the allocation of import spending across countries. Consistent with our theory, we find that larger firms pay higher prices for their imported inputs within narrowly defined products. This is in line with Kugler and Verhoogen (2009, 2012), who show that larger firms pay higher average input prices and Manova and Zhang (2012), who show that exporters buy more expensive inputs from abroad. Importantly, we find that countries differ in their ability to produce high quality inputs, with richer countries being particularly productive in producing high quality goods. This finding is reminiscent but different to Khandelwal (2010). He uses data on import prices and aggregate import shares to estimate country-level exogenous quality. We use firm-variation to estimate the slope of the quality production function within sourcing countries.

The structure of the paper is as follows. Section 2 lays out the baseline model of importing and derives the homotheticity property of import demand. Section 3 contains our test of this homotheticity property. In Section 4, we introduce a tractable extension of the baseline model and show that it can rationalize our findings. Section 5 concludes.

# 2 A Baseline Model of Importing

In this section, we present a general framework of importing. Firms' import decisions are the solution to a static profit maximization problem, there is imperfect substitutability between inputs, production is subject to constant returns and productivity is factor neutral. This framework is common in the literature and nests most contributions, e.g., Halpern et al. (2015), Gopinath and Neiman (2014), Antràs et al. (2017), Kasahara and Rodrigue (2008), Lu et al. (2016), Amiti et al. (2014) and Goldberg et al. (2010).

## 2.1 The Environment

The economy is populated by a mass of firms, indexed by $i$, which purchase multiple inputs to produce. Inputs can either be sourced domestically or can be imported. We make the usual distinction between input products and varieties. There is a set of products $\mathcal{K}$ with $n$ elements and a set of countries $\mathcal{C}$ from which the products can be sourced. Varieties are differentiated by their country of origin within the same product class. The difference between products and varieties is embedded in the technology. In particular, we assume that the production function of firm $i$ is given by

[3] In Blaum et al. (2018), we derive a sufficient statistic for the change in consumer prices resulting from a shock to the trade environment. This result does not rely on the import demand system being homothetic.

$$y = y\left(l, x; \varphi_i\right), \tag{1}$$

where $l$ denotes primary factors, $\varphi_i$ denotes firm efficiency and $x$ is an aggregator of product bundles $x_k$ defined by[4]

$$x = f\left(x_1, .., x_n; \varphi_i\right) \tag{2}$$

$$x_k = g_k\left(\left\{\eta_{ck} z_{ck}\right\}_{c \in \Sigma_{ki}}; \varphi_i\right). \tag{3}$$

Here $\eta_{ck}$ parametrizes the quality of product $k$ supplied by country $c$, and $z_{ck}$ denotes the quantity of product $k$ sourced from country $c$. Importantly, $\Sigma_{ki} \subseteq \mathcal{C}$ is the set of countries from which firm $i$ sources product $k$ (which may include the domestic country). We define the firm's *sourcing strategy* $\Sigma_i$ as the collection of sets of countries from which each product is sourced, i.e., $\Sigma_i \equiv \{\Sigma_{ki}\}_k$. For notational simplicity, we refer to the number of countries firm $i$ sources product $k$ from as $n_{ki} = |\Sigma_{ki}|$.[5]

We restrict the analysis to production functions $f$ and $\{g_k\}_k$ with constant returns to scale *given the sourcing strategy* $\Sigma_i$.[6] This restriction, which is pervasive in the literature, allows us to study the optimal choice of input quantities from a given set of countries without taking a stand on how such set is selected. Hence, we do not need to impose any restrictions on how the sourcing strategy $\Sigma_i$ is determined, e.g., whether foreign sourcing is limited by the presence of fixed costs or whether firms find their suppliers through a process of costly search. We exploit this separability between the intensive and extensive margin of firms' import demand heavily in our empirical test.

In terms of heterogeneity, we allow firms to differ in their efficiency $\varphi_i$ and their sourcing strategy $\Sigma_i$. The heterogeneity in $\Sigma_i$ reflects both firms' choices and any underlying heterogeneity in additional primitives. This amounts to allowing for unrestricted heterogeneity in any unmodeled determinant of the sourcing strategy. For example, in a model with fixed costs to foreign sourcing, we can allow such costs to be firm-specific. This is for example assumed in Halpern et al. (2015) or Antràs et al. (2017).

Regarding the market structure, we assume that firms are price takers in input markets, i.e., they can purchase any quantity at given input prices. We assume that these prices, denoted by $[p_{ck}]$, are common across firms and contain all variable transport costs. In contrast, we make no assumptions on the structure of output markets, i.e., on the nature of demand for final goods and how firms compete.

We now impose the restriction that is at the heart of our analysis. We consider environments where efficiency is factor neutral.

**Assumption 1.** *Firm efficiency is factor neutral across varieties. That is,*

$$g_k\left(\left\{\eta_{cki} z_{ck}\right\}_{c \in \Sigma_{ki}}; \varphi_i\right) = \varphi_i \hat{g}_k\left(\left\{\eta_{cki} z_{ck}\right\}_{c \in \Sigma_{ki}}\right) \tag{4}$$

[4]With a slight abuse of notation, we define the production function $f$ over the entire set $\mathcal{K}$ and can always set $\eta_{ck} = 0$ if a particular product is technologically useless.

[5]An assumption encapsulated in the nesting structure of (1)-(3) concerns the degree of substitutability between different varieties. In particular, the relative marginal product of two varieties within a product $k$ does neither depend on the production functions $q$ or $f$ nor on any allocations $(x_{k'})$ or technologies $(g_{k'})$ in different product classes $k' \neq k$. This feature will be convenient in the empirical analysis to test the theory at the product level.

[6]Controlling for the sourcing strategy is an important requirement. Consider for example the canonical model of importing with fixed costs. In this case, there are increasing returns to scale in production. However, holding the sourcing strategy fixed, the technology features constant returns to scale.

*where* $\{\hat{g}_k\}_k$ *are constant returns to scale production functions.*

We refer to the framework laid out so far, including Assumption 1, as the *baseline model of importing*. This framework is general enough to nest most of the existing theories of importing with firm heterogeneity. In particular, the literature typically considers the special case with a CES production structure of the form:

$$\hat{g}_k(.) = \left( (\eta_{Dk} z_{Dk})^{\frac{\varepsilon-1}{\varepsilon}} + (m_{Fk})^{\frac{\varepsilon-1}{\varepsilon}} \right)^{\frac{\varepsilon}{\varepsilon-1}} \text{ and } m_{Fk} = \left( \sum_{c \in \Sigma_{ki}} (\eta_{ck} z_{ck})^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}, \tag{5}$$

where $m_{Fk}$ is a bundle of foreign varieties and $D$ denotes the domestic input. (5) is a special case of (4). In addition, virtually all contributions in the literature assume that (1)-(2) are Cobb-Douglas where efficiency is by construction factor neutral and rely on fixed costs as a theory of the extensive margin. Most of our analysis does not rely on either of these specific assumptions in any way.

## 2.2 The Homotheticity of Import Demand

We now characterize firms' import demand at the intensive margin. Conditional on its sourcing strategy $\Sigma_i$ and scale $x$, the firm chooses the quantities of material inputs by solving the following cost-minimization problem:

$$\Gamma(\Sigma_i, x, \varphi_i) \equiv \min_{\{z_{ck}\}} \left\{ \sum_{\{c \in \Sigma_{ki}\}_k} p_{ck} z_{ck} \text{ s.t. } f(x_1, .., x_n; \varphi_i) \geq x \right\}, \tag{6}$$

where $x_k = \varphi_i \hat{g}_k(\cdot)$ and we denote by $z_{ck}(\Sigma_i, x, \varphi_i)$ the optimal amount sourced from country $c$ of product $k$ for a firm producing $x$ units of the input composite with efficiency $\varphi_i$ and sourcing strategy $\Sigma_i$. Because we are interested in the allocation of expenditure across countries, we focus on the firm's expenditure shares given by

$$s_{ck}(\Sigma_i, x, \varphi_i) \equiv \frac{p_{ck} z_{ck}(\Sigma_i, x, \varphi_i)}{\sum_{j \in \Sigma_{ki}} p_{jk} z_{jk}(\Sigma_i, x, \varphi_i)}.$$

The following proposition contains the central homotheticity property which is the focus of this paper.

**Proposition 1.** *The optimal expenditure shares of firm* $i$ *satisfy*

$$s_{ck}(\Sigma_{ki}, x, \varphi_i) = s_{ck}\left( [\eta_{ck}, p_{ck}]_{c \in \Sigma_{ki}} \right), \tag{7}$$

*that is, they depend only on the qualities and prices of the varieties sourced within product* $k$*. Conditional on* $\Sigma_{ki}$*, firm* $i$*'s expenditure shares are independent of productivity* $\varphi_i$*.*

*Proof.* See Section A.1 in the Appendix. □

Proposition 1 implies that, conditional on the sourcing strategy $\Sigma_{ki}$, the within-product allocation of expenditure is equalized across firms. Firm characteristics, notably firm efficiency $\varphi_i$, have no effect on the expenditure shares other than through their effect on the set of varieties $\Sigma_{ki}$. As long as firm heterogeneity is summarized by $(\varphi_i, \Sigma_i)$, this result follows directly from the assumptions of factor neutral efficiency and

constant returns to scale. The fact that only the prices and qualities of the varieties of the corresponding product $k$ matter follows from the assumption of a nested production function encapsulated in (1)-(3).

The result in Proposition 1 can be easily seen in the case of the canonical CES production function (see (5) above), where (7) reduces to the usual expression

$$s_{ck}\left(\Sigma_{ki}, x, \varphi_i\right) = \frac{\left(\eta_{ck}/p_{ck}\right)^{\rho-1}}{\sum_{j\in\Sigma_{ki}}\left(\eta_{jk}/p_{jk}\right)^{\rho-1}}. \tag{8}$$

This expression directly shows that no firm characteristic determines the pattern of spending once $\Sigma_{ki}$ is controlled for. In other words, two firms that source product $k$ from the same set of countries should have the exact same expenditure shares regardless of their efficiency $\varphi$. In Section 3 below, we test whether import demand actually satisfies this homotheticity property using French data.

**Discussion of the Extensive Margin of Importing.** We derived the homotheticity property in Proposition 1 without taking a stand on the extensive margin of importing. To do so, we relied on optimality conditions from the cost minimization problem taking the sourcing strategy as given. In general, firms select their sourcing strategy by balancing the cost reductions associated with importing from different countries with the costs of setting up relationships with foreign countries. Formally, the firm chooses its optimal size and sourcing strategy to solve:

$$\left(\Sigma^*, y^*, l^*, x^*\right) = \arg\max_{(\Sigma,y,l,x)}\left\{p\left(y\right)y - \Gamma\left(\Sigma_i, x, \varphi_i\right) - wl - \Omega\left(\Sigma\right)\right\}, \tag{9}$$

subject to (1), where $p\left(\cdot\right)$ denotes the demand that the firm faces, $\Gamma$ is the cost function defined in (6) and $\Omega(\cdot)$ provides the cost of establishing a particular sourcing strategy. The literature has typically modeled these extensive margin costs as fixed costs - see Gopinath and Neiman (2014), Halpern et al. (2015), Antràs et al. (2017) or Eslava et al. (2017). As also discussed in Antràs et al. (2017), in a context with fixed costs this framework does not necessarily predict that more efficient firms import from more countries. As long as there are complementarities across foreign varieties in the production function, the cost reduction from importing a particular variety depends on the entire sourcing strategy $\Sigma$. Thus, it might be that relatively inefficient firms source multiple varieties with low fixed costs and low quality flows while more efficient firms concentrate on few fixed cost expensive high-quality varieties. The interdependence of sourcing decisions can make the characterization of the extensive margin of importing computationally harder to solve than in the case of exports, where entry decisions are made market by market - see for example Eaton et al. (2011). We can ignore these difficulties. By conditioning on firms' sourcing strategies directly, we focus on the homotheticity of import demand at the intensive margin, which holds regardless of the mechanics of the extensive margin.

# 3 Testing for the Homotheticity of Import Demand

In this section, we assess whether firms' expenditure shares are independent of firm size as predicted by the baseline model of importing - see Proposition 1. Using data on French manufacturing firms, we show that this prediction is systematically rejected. We propose a novel theoretical mechanism that can account for this result in Section 4.

## 3.1 Data

We start by providing a general overview of the dataset. A detailed description is contained in Appendix B.1. Because we are interested in the demand for inputs, we restrict the analysis to manufacturing firms.[7] The official custom files allow us to observe import flows for every manufacturing firm in France. Between 2001 and 2006, manufacturing firms account for roughly 25% of the population of French importing firms and 50 to 55% of total import value.

| | | All Importers | | Importers of ≥2 varieties for at least 1 product | |
|---|---|---|---|---|---|
| | | Mean | Median | Mean | Median |
| Sales | | 27,529 | 3,346 | 43,378 | 6,308 |
| Workers | | 92 | 23 | 139 | 40 |
| Domestic share | | 65% | 75% | 58% | 64% |
| # imported products | | 17 | 6 | 26 | 14 |
| # sourcing countries | | 5 | 3 | 8 | 6 |
| Exporters | | 74% | 100% | 86% | 100% |
| Affiliates of corporate groups: | all | 45% | 0% | 58% | 100% |
| | with foreign subsidiaries | 24% | 0% | 33% | 0% |
| | with foreign headquarters | 13% | 0% | 18% | 0% |
| Capital intensity: | tangible assets per worker | 77 | 31 | 92 | 36 |
| Number of observations | firms × years | 187,191 | | 111,271 | |
| | firms | 49,680 | | 28,612 | |

Notes: Importing firms active in the French manufacturing industries between 2001 and 2006. Our measure of tangible assets is the book value reported in firms' balance sheets (at "historical cost"). Sales and tangible assets are expressed in thousand Euros. See Section B.1 in the Appendix for a complete description of the data.

Table 1: Characteristics of French Importers in the Manufacturing Sector

In the product dimension, import flows are classified at the 8-digit (NC8) level of aggregation, which means that the product space consists of roughly 9,300 to 9,800 products depending on the year. We define a variety as an 8-digit product coming from a particular country. Using unique firm identifiers, we match this dataset to fiscal files that contain detailed complementary accounting information. The final sample consists of an unbalanced panel of roughly 50,000 importing firms which are active between 2001 and 2006. Table 1 contains some basic descriptive results for our estimation sample. It includes a total of 187,191 firm × year observations, which implies that there are about 31,000 manufacturing importers per year. The median importing firm has 23 workers and annual sales of slightly more than 3 millions Euros. Three quarters of importing firms are also exporters, and this proportion increases to 86% for firms importing several varieties of at least one product.

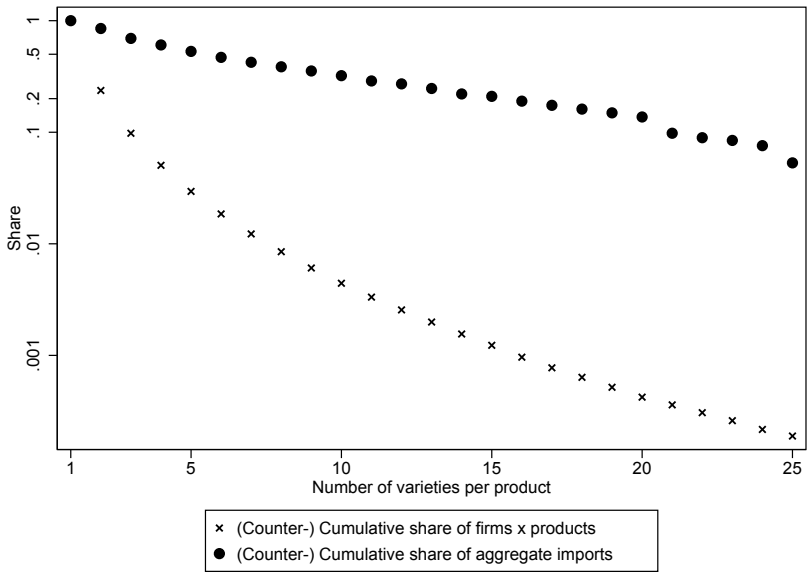
## 3.2 Sourcing Strategies

Firms' sourcing strategies are a key ingredient of the theoretical results of Section 2. Proposition 1 states that the pattern of expenditure across input varieties should be equalized among firms with a common sourcing strategy $\Sigma$. We now document two properties of firms' extensive margin of trade which underscore the importance of appropriately controlling for the sourcing strategy to implement a test of Proposition 1

[7] We focus on manufacturing firms for two reasons. First, the literature on input trade has focused on firms that are involved in the production of goods - see e.g. Halpern et al. (2015), Gopinath and Neiman (2014), Antràs et al. (2017) and Blaum et al. (2018). Second, only 5% of the firms in the service sector are importers and they account for less than 5% of economy wide imports.

empirically. We show (1) that firms, in particular large importers, routinely source a given narrowly defined product from more than a single country and (2) that the number of sourcing countries is not a sufficient statistic for firms' sourcing sets.

**Multi-Country Sourcing and Firm Heterogeneity.** Do firms source a given narrowly defined product from more than a single country? And do firms differ in the extent to which they engage in such multi-product sourcing? Figure 1 addresses these questions by displaying the distribution of the number of countries per product. Specifically, it reports the share of firm-product pairs coming from at least $C$ countries and the share of aggregate imports these firm-product pairs account for. While the majority of French firms source a limited number of varieties per 8-digit product - 90 percent of French firm-product pairs are associated with less than 3 varieties and 75 percent of them with only a single variety - there is a group of firm-product pairs that stem from many countries - about 1 percent are sourced from more than 8 countries. Importantly, these (few) firm-product pairs are very influential as they account for about 40% of aggregate imports.[8] This suggests that there is meaningful heterogeneity in firms' sourcing strategies.




Notes: This figure shows the share of firm×product pairs with at least $n$ varieties and the share of aggregate imports these firms×product pairs account for. We use 6 years of data from 2001-2006 and report the yearly average.

Figure 1: The Importance of Multi-Country Sourcing

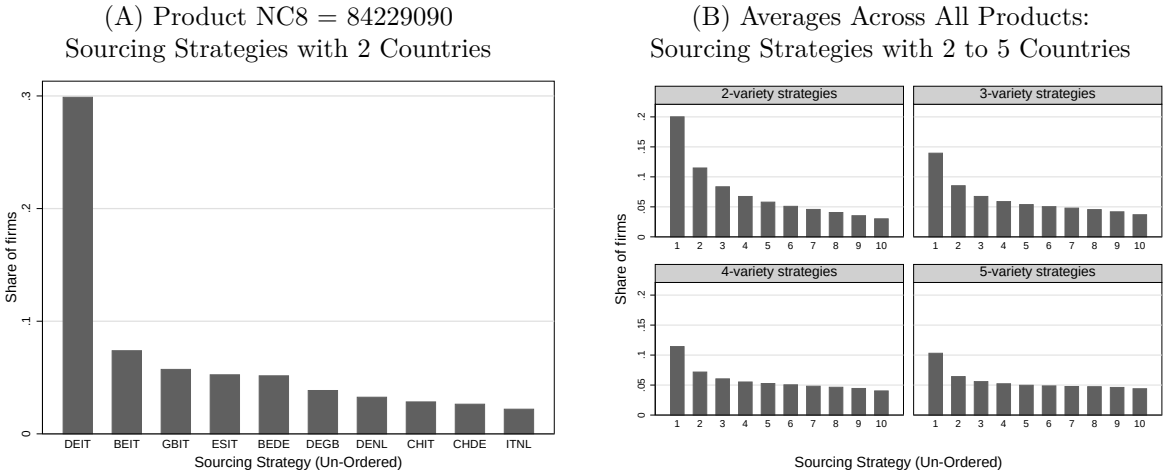
**Is the number of countries a sufficient statistic for the sourcing strategy?** Several theoretical frameworks in the literature have imposed sufficient structure on the import environment for this to be the case.[9] We show that this is not supported by the data.

Figure 2 documents the extent to which firms that agree on the number of sourcing countries *also* agree on their identity. It shows that among firms sourcing a particular product from a given number of countries,

[8] In Section B.2.1 in the Appendix, we assess the degree of concentration in spending across countries for multi-country firms. On average, the most popular variety accounts for about $50\% - 60\%$ of a firm's import budget. Thus, while there is substantial concentration, multi-country firms spend a non-trivial amount of their total spending on their non-top varieties.

[9] See e.g. Gopinath and Neiman (2014), Halpern et al. (2015), Blaum et al. (2018) or Ramanarayanan (2014)

there is substantial disagreement about the identity of these trading partners. To fix ideas, the left panel focuses on one particular product as an example: "parts of packing or wrapping machinery, including heat-shrink wrapping machinery" (NC8 84229090). We consider all firms that import this product from exactly two countries. We depict the top ten sourcing strategies (two-country tuples) ranked by popularity. If the number of countries was a sufficient statistic for the extensive margin of trade, this distribution would be degenerate. This is clearly not the case. While 30% of importers source their parts of wrapping machinery from Germany and Italy, about 20% of importers pair the Italian variety with either Belgium, Britain or Spain. Similarly, slightly more than 5% of firms have suppliers in China and supplement them with suppliers in either Germany or Italy. Thus, while Germany or Italy always appear in all top ten sourcing strategies, there is ample disagreement as to the identity of the accompanying country.




Notes: In the left panel, we show the share of firms importing product NC8 = 84229090 ("parts of packing or wrapping machinery, including heat-shrink wrapping machinery") from exactly two countries for each of the ten most popular sourcing strategies. The countries are Germany (DE), Italy (IT), the UK (GB), Spain (ES), Belgium (BE), the Netherlands (NL) and China (CH). Hence, "DEIT" refers the sourcing strategy "Germany and Italy". In the right panel, we show the share of firms by sourcing strategy when we average our data across all products for different cardinalities of the sourcing strategies (from 2-variety to 5-variety sourcing strategies). The $x$-axis denotes the rank of the considered sourcing strategy in terms of its popularity. In both panels, we use 6 years of data from 2001-2006 and report yearly averages. Only products $\times$ cardinality of sourcing strategies with more than 10 observations are considered.

Figure 2: Disagreement on the Extensive Margin

The right panel of Figure 2 shows that this pattern is not specific to the example of wrapping machinery but holds more generally. We redo the analysis of the left panel for all products and report the across-product average. In addition, we consider not only two-country strategies, but also the cases of three, four and five varieties per product.[10] We find considerable disagreement in the countries from which firms import their products, holding the number of countries in the sourcing set fixed. For example, only 20% of firms that source a given 8 digit product from exactly two countries, source this product from the most popular combination of countries. These patterns suggest that, to implement Proposition 1, we have to control for firms' entire sourcing set.

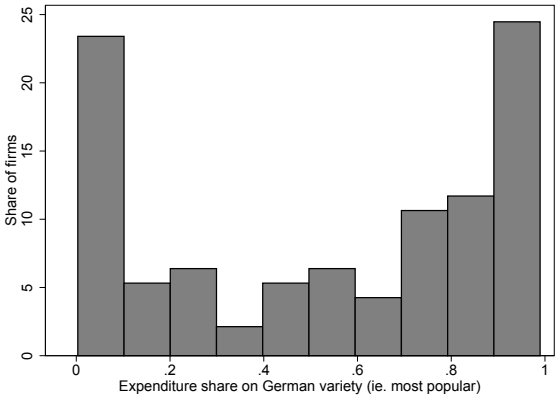
[10]Consider the figure in the upper left corner of Panel B. For each product $k$, we select all firms sourcing this product from *exactly* two countries. We then consider all distinct sourcing strategies, i.e., all combinations of two-country pairs, calculate the share of firms in each of them, rank the different sourcing strategies by their popularity and average these distribution across products. The remaining cases of three, four and five varieties per product are calculated similarly.

### 3.3 Is Import Demand Homothetic?

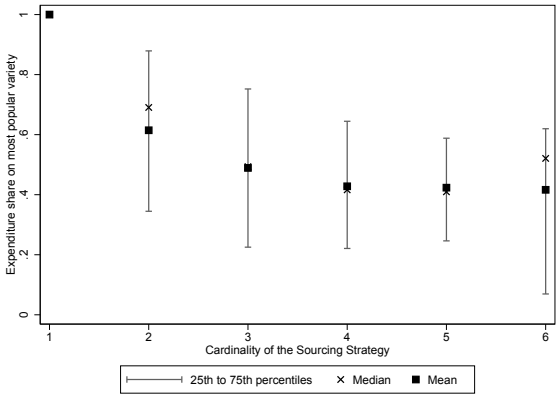
We now turn to the test of the main homotheticity prediction of the baseline model. Proposition 1 puts lots of restrictions on the data. It implies that all firms that source a particular product $k$ from a set of countries $\Sigma_k$ should have the exact same expenditure shares on each supplying country $c \in \Sigma_k$.[11] We now start to investigate to what extent this result is borne by the data. We first briefly show that there is substantial disagreement in firms' spending shares across a common set of sourcing countries. We then turn to our main analysis and show that such disagreement is not merely idiosyncratic but systematically related to firm size.

**Disagreement in Spending.** We first go back to the example considered in Figure 2 above. We consider all firms that source product NC8 84229090 ("parts of packing or wrapping machinery, including heat-shrink wrapping machinery") from the exact same two countries - Germany and Italy. There are 94 such firms in 2002.[12] In the left panel of Figure 3, we depict the cross-sectional distribution of expenditure shares on the German variety. The theory predicts that this distribution should be degenerate. Empirically, this is clearly not the case: there is substantial disagreement in how firms allocate their spending on the German vs Italian variety.

(A) Product NC8 = 84229090
Sourcing Strategy = {DE, IT}




(B) Averages Across
All Products × Sourcing Strategies




Notes: The left panel shows the distribution across firms of the share of expenditures on "parts of packing or wrapping machinery (including heat-shrink wrapping machinery)" allocated to Germany for firms sourcing this product only from Germany and Italy in 2002. The right panel depicts moments of the distribution of the expenditure share on the most popular variety within each product×sourcing strategy, depending on the cardinality of the sourcing strategy. We use 6 years of data from 2001-2006 and report the yearly averages. We restrict our sample to products×sourcing strategies with more than 5 observations.

Figure 3: Disagreement on the Intensive Margin

In the right panel of Figure 3, we consider all products and sourcing strategies and show that the pattern found in the previous example holds more generally. For each product and sourcing strategy, we define the most popular variety as the one which most firms have as their most important supplier and compute the dispersion (as measured by the interquartile range) of expenditure shares on this variety. We then average

[11] Note that we can perform the analysis at the product level because of the nesting structure in the production function (2)-(3). Because we do not observe domestic spending at the product-level, throughout this section, we focus on how foreign expenditure is allocated across the different foreign varieties. That is, the expenditure share on any particular country is computed as the ratio of expenditure in such country over total foreign expenditure.

[12] The specific product NC8 84229090 has been selected for its popularity which allows us to plot Figure 3 without violating statistical secrecy.

these dispersions across products and sourcing strategies of a given cardinality. Proposition 1 implies that there should be no dispersion across firms. This is not what we find. For example, for all products × sourcing strategies combinations involving exactly two countries, the interquartile range of the expenditure share on the most popular variety ranges from 40% to 90%.[13]

In Section B.2.2 of the Appendix, we provide additional evidence on these patterns of disagreement. In particular, we present regression evidence that a full set of sourcing strategy fixed effects still leaves ample unexplained variation of firms' spending patterns. This is inconsistent with the baseline model which implies that sourcing strategies fully determine the distribution of spending.

**Firm Size and the Concentration of Import Spending.** The disagreement on the intensive margin documented above can emerge from purely idiosyncratic forces, e.g., if particular firms have preferences for particular countries. Alternatively, more systematic forces can be at play by which the distribution of spending across sourcing countries varies across firms in a predictable way. We now provide evidence consistent with the latter alternative.

More specifically, we show that large firms have a more concentrated distribution of import spending because they allocate a larger share of their import expenditure on their top trading partner. To measure the concentration of import expenditure, we consider the Herfindahl index, the Gini index, and the share accounted for by the top varieties. We then consider the specification

$$M_{kit} = \alpha_t + \alpha_{\Sigma_{kit}} + \beta \times \ln(S_{it}) + X'_{it}\mu + u_{kit}, \tag{10}$$

where $k$,$i$,$t$ denote a product, a firm and a year, respectively, $M_{ki}$ denotes one of the concentration measures described above[14], and $\alpha_t$ are time fixed effects. Most importantly, $\alpha_{\Sigma_{kit}}$ denotes a complete set of industry × product × sourcing-strategy fixed effects. Hence, as required by the theory and by the empirical evidence on the heterogeneity in sourcing strategies documented in Figure 2 above, we control for firms' extensive margin of trade non-parametrically by comparing firms in the same industry sourcing inputs in the same product category from exactly the same set of countries. Log sales, denoted by $\ln(S_{it})$, is taken as a proxy for productivity $\varphi_i$. The vector $X_i$ contains additional firm-characteristics to control for differences in technology and the input market environment within industries.[15]

The baseline model implies that $\beta = 0$ for any moment $M_{ki}$ as firm size should not systematically affect the allocation of import spending across sourcing countries. The results are contained in Table 2. We measure products at the 8-digit level (Panel A) and at the 4th digit level (Panel B). Columns 1 and 2 consider the Herfindahl and Gini indices of the distribution of expenditure shares at the firm-product level. Columns 3-7 consider the expenditure share of the top $\tau$ varieties, where $\tau$ varies from 1 to 5.

Table 2 shows that the homotheticity property implied by Proposition 1 is soundly rejected - the estimated coefficient for sales is highly statistically significant regardless of which measure of concentration is used. Importantly, the coefficient is strongly *positive* for all specifications. Columns 1 and 2 therefore imply that larger firms feature more concentrated distributions of foreign spending holding the set of countries constant.

[13] For ease of readability, we also include the case of a single variety importers, where the expenditure share on the most popular variety is by construction unity.

[14] More precisely, $M_{ki}$ is one of the following measures of concentration: the Herfindahl index ($M_{ki} = \sum_{j=1}^{n_{ki}} s_{jki}^2$), the Gini index and the share of expenditure accounted for by the top $\tau$ varieties ($M_{ki}^{\tau} = \sum_{m=1}^{\tau} s_{ki}^{(m)}$), where $s_{ki}^{(m)}$ is the $m$th highest expenditure share of firm $i$ on product $k$.

[15] In particular, we control for capital intensity, export status and whether the firm has foreign affiliates.

Columns 3-7 then show that this higher concentration stems from larger expenditure shares on the *top* varieties, as evidenced by the decreasing coefficient for sales across columns 3 to 7. In particular, these results imply that the distribution of foreign spending of large firms first-order stochastically dominates that of their smaller counterparts. Quantitatively, a one standard deviation increase in sales is associated with an increase in the expenditure share on the top foreign variety of about 2.7 percentage points, i.e., from 75.6% to 78.3% in the Panel A specification.

| | ln Herf. | ln Gini | Expenditure share on... varieties | | | | |
|---|---|---|---|---|---|---|---|
| | | | Top 1 | Top 2 | Top 3 | Top 4 | Top 5 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| | | | Panel A: Products aggregated at the 8 dig. level | | | | |
| ln Sales | 0.017*** | 0.047*** | 0.010*** | 0.007*** | 0.004*** | 0.002* | 0.002 |
| | (0.001) | (0.005) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Impact of $\Delta_{sd}$ sales: | 0.045 | 0.124 | 0.027 | 0.017 | 0.009 | 0.004 | 0.003 |
| Sample mean: | -0.449 | -1.240 | 0.756 | 0.892 | 0.937 | 0.959 | 0.970 |
| Observations | 739,167 | 737,668 | 739,167 | 306,573 | 158,652 | 92,722 | 58,367 |
| Identified FE | 534,061 | 532,945 | 534,061 | 261,019 | 144,718 | 87,588 | 56,252 |
| of which singletons | 435,837 | 434,905 | 435,837 | 232,960 | 134,865 | 83,661 | 54,513 |
| | | | Panel B: Products aggregated at the 4 dig. level | | | | |
| ln Sales | 0.016*** | 0.040*** | 0.009*** | 0.005*** | 0.003*** | 0.002** | 0.001 |
| | (0.001) | (0.004) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Impact of $\Delta_{sd}$ sales: | 0.043 | 0.106 | 0.025 | 0.014 | 0.008 | 0.005 | 0.003 |
| Sample mean: | -0.474 | -1.164 | 0.746 | 0.883 | 0.931 | 0.954 | 0.966 |
| Observations | 597,628 | 596,722 | 597,628 | 283,009 | 158,801 | 97,756 | 64,058 |
| Identified FE | 404,647 | 404,074 | 404,647 | 231,751 | 141,864 | 91,213 | 61,308 |
| of which singletons | 321,900 | 321,463 | 321,900 | 202,799 | 130,588 | 86,435 | 59,141 |

Notes: Regressions are estimated at the importer × product level. The estimation samples include importers with more than two, three, four, five or six varieties depending on the specification observed from 2001 to 2006. Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Regressions in Panels A (B) include fixed effects for years, 4-digit industries interacted with 8-digit (4-digit) products and sourcing strategies (for the particular product). All regressions include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating foreign headquarters, and a measure of capital intensity (log tangible assets per worker). Observations are weighted to give each firm an equal weight independently of the number of imported products. $\Delta_{sd}$ sales denotes the marginal impact of an increase in sales by one standard deviation on the respective dependent variables. "Sample mean" refers to the sample mean of dependent variable.

Table 2: Non-homothetic Import Demand: Firm Size and the Concentration of Import Shares

To give a structural interpretation of these results, we now consider a set of specifications that follow more tightly from our theory. Under the restriction that the production functions take the CES form (see equation (8) above), a firm's relative expenditure share between varieties $c$ and $j$ is a log-linear function of the relative price-adjusted import qualities

$$\ln\left(s_{ck}/s_{jk}\right) = (\rho - 1)\left(\ln\left(\eta_{ck}/\eta_{jk}\right) - \ln\left(p_{ck}/p_{jk}\right)\right). \tag{11}$$

This suggests the following log-linear specification for relative spending on the top variety *vs.* the variety with the $m^{\text{th}}$ highest spending share

$$\ln\left(s_{kit}^{(1)}/s_{kit}^{(m)}\right) = \alpha_S + \alpha_t + \alpha_k + \beta \times \ln(S_{it}) + X_{it}'\mu + u_{ikt}, \tag{12}$$

where as above $\alpha_S$, $\alpha_t$ and $\alpha_k$ are sector, year and product fixed effects and $\ln(S_{it})$ denotes log sales. As the dependent variable in (12) is one particular moment $M_{ki}$, the baseline model again implies that

$\beta = 0$. Furthermore, the CES specification delivers two additional insights. First, it provides a structural interpretation for the sign of the coefficient $\beta$: as relative expenditure shares are proportional to relative price-adjusted qualities, $\beta$ can be interpreted as reflecting the correlation between firm productivity (and hence size) and price-adjusted qualities. Second, the CES specification in (11) implies that the sourcing strategy $\Sigma_{ki}$ does not need to be controlled for when *relative* expenditure shares are considered.

The results for the top five varieties of firms' sourcing sets are contained in Table 3. We also report a subset of the results when products are defined at the 4-digit level.[16] In Panel A we focus on the specification when we do not control for the sourcing strategy - as implied by the CES production function. The results are consistent with the findings in Table 2. In particular, larger firms bias their expenditure towards their most important foreign varieties. This is consistent with the above finding that the distribution of foreign spending of large firms first-order stochastically dominates that of smaller firms. For completeness, in Panel B we also report the results when we estimate (12) but control for firms' sourcing strategies using sourcing-strategy specific fixed effects as in Table 2. As before, there is a strong positive correlation between firms productivity (as proxied by sales) and the relative expenditure share on the top input variety.

Equation (11) highlights why this particular departure from the baseline model is economically meaningful. If expenditure shares are increasing in price-adjusted input quality, these results suggest that large firms bias their expenditure towards varieties of high quality. In Section 4, we indeed show that a model with a complementarity between firm productivity and input quality can rationalize these findings.

**Robustness.** In the Appendix, we perform a number of additional exercises which confirm the findings of Tables 2 and 3. In Section B.2.3, we show that the results are robust to using the number of sourcing countries as a control for the sourcing strategy. In Section B.2.4, we show that the obtained results are quantitatively very similar with employment or TFP as alternative proxies for firm efficiency. We also show that the results are robust to controlling for firm fixed effects. In Section B.2.5, we re-estimate the specification in (12) on samples of firm-product pairs that perfectly agree on the sourcing strategy. That is, for each product, we compare firms that share the exact same sourcing strategy.[17] We also estimate a specification where the coefficient on sales, $\beta$, is allowed to vary by sourcing strategy. We estimate about 6,700 coefficients and find that the majority of them are positive. In Section B.2.5, we furthermore outline an approach that allows us to estimate the coefficients in Table 3 on a substantially larger sample, because we can pool firms whose sourcing strategies coincide on at least a particular pair of countries. In all these exercises, the homotheticity property of the baseline model of importing is systematically rejected. In all specifications, we find that larger firms spend relatively more on their top varieties.

**Domestic vs. Foreign Expenditure.** The analysis so far dealt with the allocation of expenditure across foreign varieties. We now focus on the trade-off between foreign and domestic varieties. This trade-off is of special interest because it contains important information about the normative consequences of input trade. In particular, as shown in our earlier work in Blaum et al. (2018), the fraction of total material spending allocated to the domestic variety summarizes the effect of input trade on the firm's production cost in a class of firm-based models which is nested in the framework of Section 2. Additionally, because the joint distribution of firm size and domestic expenditure shares fully summarizes the effect of input trade on consumer prices,

[16] For brevity, we report only two specifications with products at the 4-digit level. The results for the remaining specifications are available upon request.

[17] This is in contrast to the approach in Table 3 where the coefficients on firm sales are identified by pooling firms with different sourcing strategies together and controlling for sourcing strategy fixed effects.

| Aggregation level: | 8-dig. products | | | | 4-dig. products | |
|---|---|---|---|---|---|---|
| | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(4)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(5)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Panel A: Not controlling for the sourcing strategy | | | | | |
| ln Sales | 0.041*** | 0.037*** | 0.033*** | 0.006 | 0.024*** | 0.011** |
| | (0.003) | (0.005) | (0.008) | (0.010) | (0.003) | (0.005) |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | | | |
| - On dependent variable | 0.109 | 0.093 | 0.081 | 0.013 | 0.063 | 0.029 |
| - On $s^{(k)}$ | -0.020 | -0.007 | -0.003 | 0.000 | -0.015 | -0.003 |
| Sample average of $s^{(k)}$ | 0.206 | 0.080 | 0.041 | 0.024 | 0.204 | 0.081 |
| Observations | 739,167 | 306,573 | 158,652 | 92,722 | 597,628 | 283,009 |
| Identified FE | 119,862 | 55,686 | 31,303 | 19,660 | 48,392 | 26,304 |
| of which singletons | 46,765 | 21,483 | 11,881 | 7,431 | 14,171 | 7,559 |
| | Panel B: Controlling for the sourcing strategy | | | | | |
| ln Sales | 0.119*** | 0.178*** | 0.202*** | 0.179*** | 0.115*** | 0.166*** |
| | (0.008) | (0.019) | (0.039) | (0.064) | (0.007) | (0.015) |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | | | |
| - On dependent variable | 0.319 | 0.451 | 0.491 | 0.418 | 0.309 | 0.427 |
| - On $s^{(k)}$ | -0.045 | -0.026 | -0.015 | -0.008 | -0.043 | -0.024 |
| Sample average of $s^{(k)}$ | 0.206 | 0.080 | 0.041 | 0.024 | 0.204 | 0.081 |
| Observations | 739,167 | 306,573 | 158,652 | 92,722 | 597 628 | 283 009 |
| Identified FE | 534,061 | 261,019 | 144,718 | 87,588 | 404,647 | 231,751 |
| of which singletons | 435,837 | 232,960 | 134,865 | 83,661 | 321,900 | 202,799 |

Notes: Regressions are estimated at the importer × product level. The estimation samples include importers with more than two, three, four or five varieties depending on the specification, observed from 2001 to 2006. Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Regressions include fixed effects for years, 4-digit industries interacted with 8-digit products in columns (1) to (4) and 4-digit products in columns (5) and (6). Regressions in panel B include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product). All regressions also include the following controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). Observations are weighted to give each firm an equal weight, independently of the number of imported products. The marginal impacts of an increase of sales by one standard deviation are computed at the mean of each sample in terms of sales, $s^{(1)}$, $s^{(2)}$, $s^{(3)}$, $s^{(4)}$ and $s^{(5)}$.

Table 3: Non-homothetic Import Demand: Firm Size and Relative Import Shares

non-homotheticities in the allocation of domestic spending can be important for the aggregate gains from input trade.

To test for such non-homotheticities, we consider the following specification for the domestic expenditure share $s_{Dit}$:

$$\ln\left(\frac{s_{Dit}}{1-s_{Dit}}\right) = \alpha_t + \alpha_{\Sigma_{it}} + \beta \times \ln(S_{it}) + X_{it}'\mu + u_{it}, \tag{13}$$

where $\alpha_\Sigma$ controls for the sourcing strategy at the firm level and all other variables are defined as above. If import demand was homothetic, the parameter $\beta$ should be zero. Because domestic shares are only observed at the firm level, we estimate (13) at the firm and not at the firm-product level.[18] The results are reported in Table 17 in Section B.2.6 of the Appendix.

Once again, we find evidence of a non-homothetic import demand. Holding the sourcing strategy fixed, we find that larger firms spend a *higher* fraction of their resources domestically (i.e., the estimated $\beta$ is positive and significant). Consistent with our previous findings for imported varieties, this result implies that large firms feature a more concentrated distribution of spending, as the domestic variety is the top variety for the vast majority of firms. Because firms' domestic shares fully summarize the productivity consequences of importing, this implies that large firms benefit less from input trade than standard models would imply.

This relationship is in sharp contrast to the effect that firm size has on the domestic share through the extensive margin of trade. Most theories of importing, where import participation is limited by fixed costs, predict a positive relation between firm size and the number of countries from which the firm sources its inputs.[19] Additionally, an increase in the number of countries in the sourcing strategy leads to a decrease in the domestic share, as love-of-variety makes importing more attractive. Hence, this extensive-margin mechanism implies a *negative* relation between firm size and the domestic share. This is exactly what we find (last column in Table 17 of the Appendix). In contrast, holding the extensive margin of trade constant, sales and domestic shares are *positively* related. This constitutes evidence of non-homotheticities in the intensive margin of trade operating in the opposite direction of those in the extensive margin. The overall relation between firm size and domestic spending is therefore the combination of these counteracting forces. In fact, we find in Blaum et al. (2018) that this unconditional correlation is essentially zero for the sample of importers.

# 4 A Model of Importing with Input Quality Choice

The findings of Section 3 are inconsistent with the baseline model of importing of Section 2 where productivity is factor neutral. In contrast, they suggest that price-adjusted input qualities have an important firm-component, which varies systematically with firm size. In this section, we provide a theory that rationalizes these findings.

[18] Proposition 1 implies that the within-product domestic expenditure share should not vary with firm size once the set of trading partners within product, $\Sigma_{ki}$, is controlled for. However, it is easy to show that Proposition 1 holds at the firm level. That is, the sourcing strategy across all products, $\Sigma_i = [\Sigma_{ik}]_k$, turns out to be a sufficient statistic for the firm-level domestic expenditure share. In particular, firms of different size should spend the same share on domestic inputs once the sourcing strategy is controlled for. More generally, in Section O.3 of the Online Appendix, we also show that there is plenty of residual variation in domestic spending shares once the sourcing set is controlled for. When controlling for the sourcing strategy with a set of interactions of product and country fixed effects, more than one third of the cross-sectional variation in domestic shares is unexplained.

[19] This prediction has been extensively corroborated in the data. See for example Halpern et al. (2015) for Hungary, Gopinath and Neiman (2014) for Argentina, Blaum et al. (2018) for France, Antràs et al. (2017) for the US, or Eslava et al. (2017) for Colombia.

## 4.1 Theory

We start from the framework laid out in Section 2. In contrast to the baseline model, where all countries offer a single homogeneous input, we now assume that each country offers its input in a range of qualities and that firms choose from this quality schedule. In addition, input quality $q$ and firm productivity $\varphi$ are assumed to be complements in production. In particular, the service flow $\eta$ of a physical unit of input of quality $q$ for a firm with efficiency $\varphi$ is given by

$$\eta(q,\varphi) = \left(\gamma\varphi^{\theta} + (1-\gamma)\, q^{\theta}\right)^{1/\theta}. \tag{14}$$

If $\gamma = 0$, input service flows are only driven by the quality of inputs as in the baseline model. The parameter $\theta$ governs the elasticity of substitution between input quality $q$ and firm efficiency $\varphi$.

We specialize our theory to the case where firms combine the set of inputs from different countries in a CES way as in (5) and use a single product from potentially multiple countries.[20] Hence, the production structure is still given by (1)-(3), where $x_k$ in (3) now reads

$$x = \left(\sum_{c=1}^{\Sigma_i}\left(\int_{q\in\mathcal{Q}_c}\eta(q,\varphi)\, z_c(q)\, dq\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}, \tag{15}$$

where $\mathcal{Q}_c = \left[\underline{q},\infty\right)$ is the set of available qualities in country $c$, $\eta(q,\varphi)$ is the firm-specific input service flow given in (14), $z_c(q)$ is the quantity of quality-$q$ inputs sourced from country $c$ and $\rho > 1$ is the elasticity of substitution across sourcing countries. Note that inputs of different qualities within countries are considered to be perfect substitutes. Thus, generically, firms will buy a single quality vintage within each sourcing country. Importantly, if quality and productivity are complements, firms with different productivity will optimally select different input qualities from a given sourcing country.

On the supply side, we assume that countries differ in their ability to produce quality. In particular, country $c$ can produce inputs of quality $q$ at marginal cost

$$MC_c(q) = \frac{1}{A_c} q^{\beta_c}. \tag{16}$$

Here $A_c$ parametrizes the quality-neutral efficiency of country $c$ and $0 < \beta_c < 1$ governs the elasticity of marginal costs with respect to quality of country $c$.[21] Hence, there are constant returns to scale in production for a given quality $q$, but producing higher quality inputs is costly. Under the assumption of perfectly competitive input markets, the price of quality-$q$ inputs from country $c$ is given by $p_c(q) = \frac{1}{A_c} q^{\beta_c}$. If $\beta_c$ is low, the elasticity of prices with respect to quality in country $c$ is small. We therefore refer to low-$\beta$ countries as countries with a comparative advantage in providing high quality inputs. We assume that in principle all countries can produce every quality in $\mathcal{Q}_c$. However, if $\beta_c$ is large, high quality inputs may be prohibitively expensive.

[20] The latter restriction is only for notational simplicity. In the empirical analysis below, we again only use the within-product variation by always including product fixed effects.

[21] The restriction that $0 < \beta < 1$ is a necessary condition for the input service flow per unit of spending $\eta/MC$ to be non-monotone in $q$, so that the optimal quality choice is interior. In this case, firms of different productivity will choose different qualities.

**Optimal Quality Choice Within Sourcing Countries.**

Consider a firm sourcing from country $c$. This importing firm has access to a quality-price *schedule* $(q, p_c(q))$. Because inputs of different quality of the same sourcing county are perfect substitutes in production, the optimal quality choice maximizes the input service flow per unit of import spending

$$\xi_c(\varphi) \equiv \max_{[z_c(q)]} \left\{ \int_q \eta(q,\varphi) z_c(q)\, dq \text{ s.t. } \int_q p_c(q) z_c(q)\, dq = 1 \right\}, \tag{17}$$

where $\eta(q,\varphi)$ is given by (14) and $p_c(q) = \frac{1}{A_c} q^{\beta_c}$. While the firm can in principle buy inputs of multiple qualities, the linearity of the problem implies that only a single quality is optimally chosen. The solution to (17) is contained in the following proposition.

**Proposition 2.** *Let $\theta < 0$. Then* (17) *has a unique and interior solution. The optimal input service flow per unit of spending $\xi_c(\varphi)$ is given by*

$$\xi_c(\varphi) = \gamma^{1/\theta} \varphi^{1-\beta_c} \left( \frac{1}{1-\beta_c} \right)^{1/\theta} \left( \frac{\beta_c}{1-\beta_c} \frac{\gamma}{1-\gamma} \right)^{-\frac{1}{\theta}\beta_c} A_c. \tag{18}$$

*Moreover, the optimal quality choice $q_c(\varphi)$, the import service flow $\eta(q_c(\varphi),\varphi) \equiv \eta_c(\varphi)$ and the price paid $p_c(q_c(\varphi)) \equiv p_c(\varphi)$ are*

$$q_c(\varphi) = \left( \frac{\beta_c}{1-\beta_c} \frac{\gamma}{1-\gamma} \right)^{1/\theta} \varphi \tag{19}$$

$$\eta_c(\varphi) = \left( \frac{\gamma}{1-\beta_c} \right)^{1/\theta} \varphi \tag{20}$$

$$p_c(\varphi) = \frac{1}{A_c} \left( \frac{\beta_c}{1-\beta_c} \frac{\gamma}{1-\gamma} \right)^{\frac{1}{\theta}\beta_c} \varphi^{\beta_c}. \tag{21}$$

*Proof.* See Section A.2 in the Appendix. □

Proposition 2 shows that, as long as quality and productivity are complements ($\theta < 0$), there is a unique optimal input quality for each importing firm. As seen in (19), more productive firms purchase inputs of higher quality as they perceive a higher return to quality. In this way, the within-country quality choice problem endogenously delivers input service flows $\eta_c(\varphi)$ and prices $p_c(\varphi)$ at the country level that are firm-specific. More productive firms pay a higher price for, and derive greater service flows from, inputs of country $c$. Note also in (21) that these firm-specific input prices scale with firm productivity with an elasticity determined by $\beta_c$. As shown below, this elasticity turns out to be the crucial parameter to generate a systematic relation between the concentration of expenditure across countries and firm size.

**Allocation of Spending Across Countries.**

In the model with input quality choice, firms with different productivity optimally buy different inputs within countries at different prices. The baseline model of Section 2, like most models of international trade, abstracts from such considerations and assumes that all firms buy the same input variety from a given country at the same price.

We now characterize the allocation of spending across countries in the model of quality choice. We also connect this model with the baseline model by proving a convenient equivalence result. As far as the allocation *across* countries is concerned, the model of quality choice is isomorphic to the baseline model of importing with a single input per country as in the class of Section 2 *except* that productivity is not factor neutral but biased towards countries with a comparative advantage in high quality inputs.

**Proposition 3.** *Let $\theta < 0$. The model with input quality choice implies that firm $i$'s relative expenditure share between country $c$ and $j$ is given by*

$$\frac{s_c(\varphi)}{s_j(\varphi)} = \varphi^{-(\rho-1)(\beta_c-\beta_j)} \left(\frac{Z_c}{Z_j}\right)^{\rho-1}, \tag{22}$$

*where*

$$Z_c = A_c \left( \left(\frac{1-\gamma}{\beta_c}\right)^{\beta_c} \left(\frac{\gamma}{1-\beta_c}\right)^{1-\beta_c} \right)^{1/\theta}. \tag{23}$$

*Moreover, the model predicts the exact same firm-level expenditure shares across countries as the model of importing of Section 2 as long as the input aggregator $x_k$ in (3) is given by*

$$x\left([z_c]_{c\in\Sigma_i};\varphi\right) = \left(\sum_{c=1}^{\Sigma_i} \left(\varphi^{1-\beta_c} z_c\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}, \tag{24}$$

*and the price of inputs from country $c$ is common across firms and given by $p_c = Z_c^{-1}$.*

*Proof.* See Section A.3 in the Appendix. □

Proposition 3, especially equation (22), shows that the model with input quality choice generates an import demand system across countries that is non-homothetic: relative expenditure shares vary systematically with firm productivity. In particular, more productive firms spend relatively more on countries which have a comparative advantage in quality production, i.e., whose quality elasticity $\beta_c$ is low.[22] In this way, the model not only generates systematic departures from the homothetic demand system of the baseline model but exactly predicts a log-linear relationship between relative expenditure shares and firm-productivity as estimated in Table 3:

$$\ln\left(\frac{s_c}{s_j}\right) = -(\rho-1)(\beta_c-\beta_j) \times \ln\varphi + \delta_j + \delta_c, \tag{25}$$

where $\delta_j$ and $\delta_c$ are country-specific constants.[23]

Equation (25) highlights the importance of country heterogeneity in the production of quality. With a common quality-cost elasticity $\beta$ across countries, expenditure shares are again equalized across firms and

[22] Recall that we assume the elasticity of substitution across inputs, $\rho$, to exceed unity. If $\rho$ was smaller than one, all inputs would be essential and firms would need to import from all countries. In addition, using variation in trade costs, prices and bilateral trade flows, a vast literature in international trade has found estimates of this parameter that largely exceed unity - see Hillberry and Hummels (2013) for a survey.

[23] Equation (25) has a similar structure to the non-homothetic demand system of Fieler (2011). The economics, however, is different. In our context, non-homotheticities arise from the interaction of firms' quality choices and differences in countries' abilities to produce high quality inputs. In Fieler (2011), non-homothetic spending patterns are a consequence of non-constant substitution elasticities across goods.

the import demand system is homothetic. Hence, demand non-homotheticities arise from the *interaction* of quality-productivity complementarities at the firm-level and differences in comparative advantage across countries encapsulated in $\beta_c$.

The second part of Proposition 3 contains an equivalence result that clarifies the connection between the model with quality choice and the model of Section 2. In particular, the quality choice model is isomorphic to a model where all countries sell a homogeneous input at a common prices, as long as productivity does not enter in a factor neutral way, but rather as in equation (24). That is, the micro-founded model of quality choice delivers a country-level model where productivity is biased towards countries with a comparative advantage in quality production.

To see this, let $MP_c(\varphi) = \frac{\partial x\left([z_c]_{c\in\Sigma_i};\varphi\right)}{\partial z_c}$ denote the marginal product of inputs from country $c$ for a firm with productivity $\varphi$. Equation (24) implies that, whenever $\beta_c < \beta_j$, we have that

$$\frac{\partial}{\partial\varphi}\left(\frac{MP_c(\varphi)}{MP_j(\varphi)}\right) = -(\beta_c-\beta_j)\frac{\rho-1}{\rho}\varphi^{-\left((\beta_c-\beta_j)\frac{\rho-1}{\rho}+1\right)}\left(\frac{z_c}{z_j}\right)^{-1/\rho} > 0, \tag{26}$$

so that more productive firms perceive a higher relative marginal product of inputs from countries that are efficient at producing quality (i.e., countries with low $\beta$). In this sense, productivity is biased towards low-$\beta$ countries (see Acemoglu (2007)).[24] Note also that with a common quality-cost elasticity $\beta$ across countries, firm productivity enters the production function in a factor neutral way and the model's cross-country implications are isomorphic to the baseline model even though input quality and prices within sourcing countries vary systematically across importing firms.

## 4.2 Bringing the Model to the Data

We now provide empirical evidence for the novel implications of the framework outlined above and show that this framework can rationalize the evidence on non-homotheticities we presented in Section 3. We proceed in three steps. First, we exploit the information contained in firms' input prices to estimate the quality-cost elasticities $\beta_c$ . We find that more productive firms pay higher prices for a given product stemming from a particular country. Importantly, as highlighted in Proposition 2, the degree to which this is the case allows us to identify $\beta_c$. Second, given these estimates, we turn to the analysis of the allocation of expenditure across countries. Consistent with Proposition 3 and equation (25), we show that productive firms indeed have relatively high expenditure shares on low-$\beta$ sourcing countries. Finally, we show that low-$\beta$ countries indeed have high *average* expenditure shares. This implies that larger firms have a more concentrated distribution of expenditure on their top suppliers, consistent with the empirical evidence of Section 3.

**Step 1: Firm Size, Input Prices and the Quality Elasticity $\beta_c$.**

Proposition 2, in particular (21), implies that input prices from country $c$ and firm productivity are related by

[24] Consider a production function $F(x_1,...,x_J;A)$ where $A$ is a technology index. Then $A$ is biased towards factor $i$ relative to factor $j$ if and only if

$$\frac{\partial}{\partial A}\left[\frac{\partial F(x;A)/\partial x_i}{\partial F(x;A)/\partial x_j}\right] \geq 0.$$

$$\ln p_{ci} = \alpha_c + \beta_c \times \ln \varphi_i, \tag{27}$$

i.e., import prices vary positively with the efficiency of the importing firm and the elasticity is governed by $\beta_c$. We first consider the basic relationship between prices and firm size. To do so, we restrict $\beta$ to be constant across countries and consider the specification

$$\ln (p_{ickt}) = \alpha_k + \alpha_c + \alpha_t + \beta \times \ln (S_{it}) + \alpha_S + X'_{it}\mu + u_{ickt}, \tag{28}$$

where $p_{ickt}$ is the price (unit value) firm $i$ pays for product $k$ inputs from country $c$ in year $t$ and $\ln (S_{it})$ denotes log-sales as previously.[25] Note that (28) contains country ($\alpha_c$) and product ($\alpha_k$) fixed effects so that we compare variety-specific prices paid by firms of different size. As part of the firm-specific controls $X_{it}$, we also include the size of the shipment, measured by a quantity index at the firm-country-product-year level, to ensure that bigger firms do not pay different prices simply because they order in bulk. Finally, we also control for sector fixed effects ($\alpha_S$) and other firm-level characteristics to isolate the effects of firm size (or productivity) on import prices. This specification is akin to the one in Kugler and Verhoogen (2012), who use Colombian data to document that larger plants pay more for their inputs.[26]

The results are contained in Table 4. Consistent with Proposition 2, larger firms indeed pay higher prices for the same 8-digit product coming from a particular country. Column 1 establishes this fact with industry, product and country fixed effects. Column 2 adds a control for shipment size. We see that large shipments trade at lower prices on average, which may reflect bulk discounts. Because larger firms have on average larger shipments, the estimated elasticity of import prices with respect to firm sales increases to 7.3%. In columns 3 and 4, we show that this elasticity is unchanged when we control for a whole set of industry-product-country interactions and when we condition on additional firm characteristics.[27]

In the analysis of Table 4, we restricted $\beta$ to be constant across countries. As highlighted in Proposition 3, to generate a non-homothetic import demand system across countries, the quality elasticities need to vary across countries - see (25). We therefore now estimate (28) allowing $\beta$ to vary by country. More precisely, we allow $\beta$ to vary at the country-product level and hence estimate cross-country differences in the quality elasticity for different products[28]:

$$\ln p_{ickt} = \alpha_{ckt} + \beta_{ck} \ln S_{it} + \alpha_{Sck} + X'_{it}\mu_{ck} + u_{ickt}. \tag{29}$$

Here $\beta_{ck}$ denotes the quality elasticity estimated at the country-product level and $\alpha_{ckt}$ denotes a product-

[25] We use $\beta$ as the notation for the coefficient in the empirical specification. This parameter would coincide with our structural quality elasticity $\beta$ if we were to use physical productivity $\varphi_i$ instead of firm sales. Under our assumption that sales and productivity are positively related, our empirical estimate is a monotone transformation of the structural quality elasticity. Because the theoretical implication in (25) only uses the *ranking* of $\beta_c$, we for simplicity refer to the empirical estimate $\beta$ as a direct estimate of the quality elasticity.

[26] Kugler and Verhoogen (2012) measure average input prices, including domestic and imported materials, without making a distinction between them. Kugler and Verhoogen (2009) distinguish domestic and imported input prices and establish that importers pay more for either type. However, they do not condition on the particular country source of the input, nor on the firm sourcing strategy, as we do. Additionally, they do not look at firm size but rather focus on import status.

[27] Online Appendix O.4 shows that our results are robust to using employment or TFP as alternative empirical proxies of firm productivity.

[28] Allowing for the quality elasticity to vary across products within countries is important. In specifications where this parameter is constrained to be common across all products within a given country (i.e. $\beta_c = \beta_{ck}$), our results are in general weaker and sometimes reversed. This is due to the fact that such "pooled" specifications do not take into account that many products are only sourced from subsets of countries (e.g., textile is not imported from Germany) and that countries often have a comparative advantage in particular products. Our preferred specification, which is estimated at the product level, generically addresses these issues.

| Dependent variable: | $\ln p_{ickt}$ | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| ln Sales | 0.024*** | 0.073*** | 0.080*** | 0.077*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| ln Shipment Quantity | | -0.224*** | -0.211*** | -0.211*** |
| | | (0.001) | (0.001) | (0.001) |
| Other controls: | No | No | No | Yes |
| Fixed effects: | Years | Years | Years | Years |
| | Industry | Industry | Industry x | Industry x |
| | Products | Products | Products x | Products x |
| | Countries | Countries | Countries | Countries |
| Impact of a 1sd increase in sales (from mean): | | | | |
| | 0.063 | 0.194 | 0.213 | 0.206 |
| Observations | 2,195,456 | 2,195,456 | 2,195,456 | 2,141,864 |
| Identified FE | 10,998 | 10,998 | 620,341 | 611,882 |
| of which singletons | 3 | 3 | 309,873 | 306,652 |

Notes: Regressions at the importer × product × country level, importers of at least 2 varieties of the considered product. These importers are observed from 2001 to 2006. Robust standard errors in parentheses with ***,** and * respectively denoting significance at the 1%, 5% and 10% levels. "Other controls" denote a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating that the group has foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). The sets of fixed effects that are inserted in each specification are indicated in each column. Observations are weighted to give each firm an equal weight.

Table 4: Import Prices and Firm Size

country-year fixed effect.

In our main specification, to be consistent with our analysis in Section 3, we define products at the 8-digit level. However, we also consider a version where we estimate $\beta_{ck}$ at the 4-digit level. We do so to increase sample size as many country-product pairs contain few firms when products are defined at the 8-digit level. Because the units of the different 8-digit products are not comparable, we do not simply aggregate the data. Instead, we estimate (29) at the 8-digit level restricting $\beta_{ck}$ to be common within 4-digit product categories.

Table 5 reports summary statistics of the distribution of estimated quality elasticities. To increase statistical power, we restrict our analysis to product-country pairs that are imported by at least 10 firms and we report our results for different choices of this cutoff. For the fine specification, our procedure yields more than 30,000 estimated coefficients in the sample of country-products with at least 10 importers. For the coarse specification, we estimate more than 15,000 elasticities. Naturally, these numbers decline as we increase the cutoff of importing firms to 50 or 100 importers.

Our theory requires the quality elasticity $\beta$ to be positive for the firms' problem to be well-defined.[29] Table 5 shows that country-product pairs with statistically significant negative elasticities (at the one percent level) are rare and account for tiny amounts of import volume. Consider for example the 4-digit (8-digit) classification with at least 50 importers. The 1,004 (755) elasticities that are estimated to be significantly positive account for about 30% (10%) of aggregate import volume. In contrast, the 228 (179) significantly negative estimates account for only 1% (0.5%) of aggregate imports. For our baseline analysis, we therefore focus on the sample of country-product pairs with positive and significant elasticities. However, we also show that our results extend to the population of positive (but not necessarily significant) coefficients.

In Figure 4, we display the distribution of estimated coefficients $\beta_{ck}^{8DIG}$ graphically.[30] In the left panel,

[29] If $\beta$ was negative, prices would be decreasing in quality and all firms would buy the highest quality input from a given country.

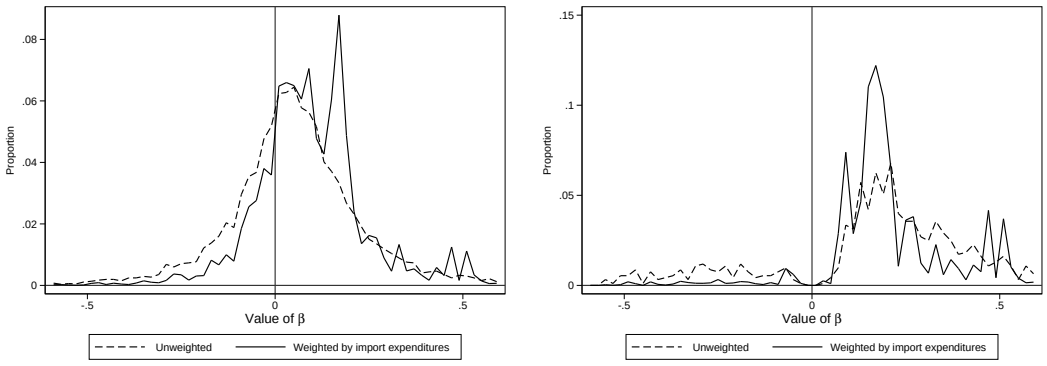
[30] Note that the "spikiness" of the distribution is a reflection of the high concentration of the trade flow data, whereby few

| Aggregation level: | 8 digit products, $\beta_{ck}^{8DIG}$ | | | | 4 digit products, $\beta_{ck}^{4DIG}$ | | | |
|---|---|---|---|---|---|---|---|---|
| | Negative betas | | Positive betas | | Negative betas | | Positive betas | |
| | All | Sig. | All | Sig. | All | Sig. | All | Sig. |
| Min obs. / variety: | $> 10$ importers | | | | $> 10$ importers | | | |
| Nb. coef. | 13,145 | 1,830 | 17,053 | 2,809 | 6,624 | 769 | 8,660 | 1,701 |
| Share coef. | 0.435 | 0.061 | 0.565 | 0.093 | 0.433 | 0.050 | 0.567 | 0.111 |
| Share imports | 0.305 | 0.042 | 0.664 | 0.216 | 0.293 | 0.031 | 0.697 | 0.321 |
| Min obs. / variety: | $> 50$ importers | | | | $> 50$ importers | | | |
| Nb. coef. | 2,854 | 179 | 5,350 | 755 | 2,875 | 228 | 5,136 | 1,004 |
| Share coef. | 0.348 | 0.022 | 0.652 | 0.092 | 0.359 | 0.028 | 0.641 | 0.125 |
| Share imports | 0.079 | 0.005 | 0.310 | 0.107 | 0.188 | 0.013 | 0.587 | 0.308 |
| Min obs. / variety: | $> 100$ importers | | | | $> 100$ importers | | | |
| Nb. coef. | 901 | 37 | 2,465 | 429 | 1,546 | 110 | 3,448 | 801 |
| Share coef. | 0.268 | 0.011 | 0.732 | 0.127 | 0.310 | 0.022 | 0.690 | 0.160 |
| Share imports | 0.038 | 0.003 | 0.225 | 0.091 | 0.159 | 0.011 | 0.456 | 0.230 |

Notes: This table complements figure 4 and describes the distribution of the product $\times$ country specific quality elasticity $\beta_{ck}$ in equation (27). The subset of coefficients that are significant at the 1 percent level are denoted by "Sig."

Table 5: Estimated Quality Elasticities $\beta_{ck}$

we consider the entire population of estimated elasticities; in the right panel, we only consider significant estimates. The solid line depicts the distribution when we weight the respective country-product cells by their aggregate import value. The dashed line is the unweighted distribution. As anticipated in Table 5, varieties with positive $\beta$'s are in the majority and - more importantly - account for most of the import value. Moreover, for the significant estimates, the entire mass of import value is concentrated on positive elasticities. Importantly, Figure 4 shows that there is substantial heterogeneity in the estimated quality elasticities - both for the population of positive and the sample of significant estimates. This heterogeneity is essential to generate a non-homothetic sourcing patterns across countries.




Notes: This figure represents the histograms (by bins of 0.02) of the 8 dig. product $\times$ country specific sales coefficient $\beta_{ck}^{8DIG}$ in equation (27). In the left panel, we depict all coefficients; in the right panel, we depict only coefficients that are significant at the one percent level. The solid line weights by the import expenditures, the dashed line reports the unweighted results. Only varieties imported by more than 50 French firms (between 2001 and 2006) are considered.

Figure 4: The Distribution of Quality Elasticities $\beta_{ck}$

varieties account for a large share of aggregate imports. See Section B.2.1 in the Appendix where we discuss this in more detail and report statistics of the distribution of the value of import flows.

**Step 2: Non-Homothetic Import Demand and Differences in Comparative Advantage**

The main implication regarding the allocation of expenditure across countries is contained in Proposition 3: the elasticity of relative expenditure shares with respect to firm productivity is determined by the *differences* in the quality elasticity $\beta$. In particular, more productive firms spend relatively more on countries with a comparative advantage in the production of high quality inputs, i.e., countries where $\beta$ is small. Moreover, the relationship of relative expenditure shares and firm productivity is predicted to be log-linear. Using the estimates of $\beta_{ck}$ recovered above, we can directly test this implication by considering the specification

$$\ln(s_{crit}) - \ln(s_{jrit}) = \alpha_c + \alpha_j + \alpha_r + \alpha_t + \lambda \times \left(\beta^r_{ck} - \beta^r_{jk}\right) \ln(S_{it}) + \alpha_S + X'_{it}\mu + u_{cjrit}, \tag{30}$$

where $\alpha_c$ and $\alpha_j$ are fixed effects for countries $c$ and $j$ as implied by (25) and $r$ denotes the product classification corresponding to the estimate of $\beta$ (i.e., either at the 4 or 8-digit level). Importantly, $\left(\beta^r_{ck} - \beta^r_{jk}\right) \ln(S_{it})$ is now directly observable. Our theory implies that $\lambda < 0$.

The results are reported in Table 6. We report estimates of $\lambda$ for both the 8-digit and 4-digit product specifications as well as for the different subsamples of Table 5. Because $\beta_{ck}$ is estimated, we bootstrap the standard errors to account for the generated regressor in (30). In the first three columns of each product specification, we consider product-country pairs with a positive and significant estimated $\beta$. In columns 4 and 8, we consider all varieties with positive but not necessarily significant $\beta$'s. Doing so results in a substantial increase in sample size. For most specifications, we indeed find a negative estimate of $\lambda$ which, in most cases, is statistically significant. Hence, larger firms tend to disproportionately allocate their spending towards countries with a comparative advantage in producing high quality inputs. This is exactly what the theory implies.

| Dep. Var.: | $\ln(s_{crit}) - \ln(s_{jrit})$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Aggregation level: | 8 digit products | | | | 4 digit products | | | |
| Sample of betas: | Positive sig. | Positive sig. | Positive sig. | Positive | Positive sig. | Positive sig. | Positive sig. | Positive |
| Min obs./variety: | 10 | 50 | 100 | 50 | 10 | 50 | 100 | 50 |
| $\left(\beta_{ck} - \beta_{jk}\right)$ | -0.005*** | -0.031 | 0.040 | -0.046*** | -0.001 | -0.019*** | -0.183*** | -0.014*** |
| $\times \ln$ Sales | (0.000) | (0.022) | (0.057) | (0.011) | (0.000) | (0.004) | (0.059) | (0.002) |
| Observations | 80,618 | 67,996 | 60,743 | 724,244 | 379,424 | 367,885 | 357,487 | 1,743,461 |
| Identified FE | 984 | 526 | 446 | 1,515 | 840 | 632 | 578 | 1,166 |
| of which singletons | 30 | 13 | 15 | 2 | 17 | 4 | 3 | 3 |

Notes: Regressions at the importer × product × country pairs level, importers of at least 2 varieties of the considered product. These importers are observed from 2001 to 2006. We rely on 100 bootstrap iterations of the entire procedure (including the first stage estimation of the $\beta_{ck}$'s and the second stage regressions reported in this table) to compute the reported standard errors while taking account of the fact that that our explanatory variable of interest is a generated regressor. The subset of $\beta_{ck}$ coefficients that are significant at the 1 percent level in first stage regressions is denoted by "sig." ***,** and * denote significance at the 1%, 5% and 10% levels, respectively. All specifications include the following controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating that the group has foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). All regressions also include year fixed effects, industry fixed at the four digit level, product fixed effects and country level fixed effects. Observations are un-weighted.

Table 6: Non-Homothetic Import Demand and Comparative Advantage in Quality Production

Proposition 2 also has strong implications for how relative prices *across* countries should vary with firm-

size. In particular, (21) implies that

$$\ln\left(\frac{p_c}{p_j}\right) = (\beta_c - \beta_j) \times \ln\varphi + \delta_j + \delta_c,$$

where $\delta_j$ and $\delta_c$ are country-specific constants. Crucially, the relationship between relative prices and firm productivity should be of the *opposite* sign of the one between relative expenditure shares and firm productivity. This is exactly what we find empirically. More specifically, in Table 7 we report the results of the specification (30) where we use relative unit values $\ln(p_{crit}) - \ln(p_{jrit})$ instead of relative expenditure shares as the dependent variable. In contrast to Table 6, we now find a coefficient that is positive and significant in most specifications. Hence, larger firms have *relatively higher* expenditure shares on low-$\beta$ countries but pay *relatively lower* prices. Recall, that $\beta$ was estimated using the *within-country* relationship between prices and firm size. In contrast, Table 7 focuses on the *cross-country* relationship between relative prices and firm size.

| Dep. Var.: | $\ln(p_{crit}) - \ln(p_{jrit})$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Estimation level for $\beta_{ck}$ | 8 digit products | | | | 4 digit products | | | |
| Sample of betas: | Positive sig. | Positive sig. | Positive sig. | Positive | Positive sig. | Positive sig. | Positive sig. | Positive |
| Min obs./variety: | 10 | 50 | 100 | 50 | 10 | 50 | 100 | 50 |
| $(\beta_{ck} - \beta_{jk})$ | -0.001*** | 0.047*** | 0.066*** | 0.018*** | 0.000 | 0.003*** | 0.052*** | 0.001*** |
| $\times$ ln Sales | (0.000) | (0.008) | (0.022) | (0.003) | (0.000) | (0.001) | (0.012) | (0.001) |
| Observations | 79,039 | 66,716 | 59,813 | 706,905 | 1,902,112 | 1,863,541 | 1,830,332 | 7,644,477 |
| Identified FE | 981 | 525 | 445 | 1,515 | 4,335 | 3,492 | 3,214 | 8,828 |
| of which singletons | 28 | 12 | 14 | 2 | 329 | 143 | 105 | 241 |

Notes: Regressions at the importer × 8 or 4 dig. product × country pairs level, importers of at least 2 varieties of the considered product. These importers are observed from 2001 to 2006. We rely on 100 bootstrap iterations of the entire procedure (including the estimation of the $\beta_{ck}$'s and the second stage regressions reported in this table) to compute the reported standard errors while taking account of the fact that that our explanatory variable of interest is a generated regressor. The subset of $\beta_{ck}$ coefficients that are significant at the 1 percent level in first stage regressions is denoted by "sig." ***,** and * denote significance at the 1%, 5% and 10% levels, respectively. All specifications include the following controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating that the group has foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker) as well as (ln) shipment sizes in each product, as measured by the quantity index of the customs data. All regressions also include year fixed effects, industry fixed at the four digit level, product fixed effects and country level fixed effects. Observations are un-weighted.

Table 7: Relative Prices, Firm Size and Comparative Advantage in Quality Production

**Step 3: Are Low-$\beta$ Countries Large Sellers?**

Table 6 shows that large firms tend to bias their spending towards low-$\beta$ countries. In Section 3, we showed that large firms concentrate their spending on their most important varieties. For these findings to be internally consistent, it needs to be the case that firms spend on average a large fraction of their spending on low-$\beta$ countries. This is exactly what we find empirically.

To see this, let $s_{ikt}^{[j]}$ be the expenditure share of firm $i$ on the country with the $j$'th lowest quality elasticity $\beta$ within product $k$. Let the *relative ranked* expenditure share be

$$m_{ikt}^{jl} \equiv s_{ikt}^{[j]} / s_{ikt}^{[l]} \text{ where } j < l.$$

If firms spend relatively more on low-$\beta$ countries, we expect that $m_{ikt}^{jl} > 1$. In Table 8, we focus on the sample of positive and significant $\beta$'s and report quantiles of the distribution of $m_{ikt}^{jl}$ as well as the fraction

of firm-product pairs for which $m_{ikt}^{jl} < 1$.

| Nb of obs. | $10^{\text{th}}$ perc. | $25^{\text{th}}$ perc. | $50^{\text{th}}$ perc. | $75^{\text{th}}$ perc. | $90^{\text{th}}$ perc. | Share of $m_{ik}^{jl} < 1$ | Min obs. / variety |
|---|---|---|---|---|---|---|---|
| Aggregation level: 8 digit products ($m_{ik}^{jl,8DIG}$) | | | | | | | |
| 1,787 | 0.8 | 4.3 | 23.9 | 126.8 | 545.1 | 0.112 | >10 |
| 569 | 5.7 | 14.7 | 55.6 | 207.3 | 618.9 | 0.014 | > 50 |
| 368 | 8.5 | 23.1 | 66.9 | 229.7 | 687.1 | 0.008 | > 100 |
| Aggregation level: 4 digit products ($m_{ik}^{jl,4DIG}$) | | | | | | | |
| 3,157 | 1.5 | 9.6 | 50.1 | 204.5 | 782.3 | 0.077 | > 10 |
| 1,806 | 7.1 | 24.0 | 81.4 | 281.3 | 877.8 | 0.018 | > 50 |
| 1,386 | 11.8 | 30.1 | 90.6 | 283.7 | 816.8 | 0.006 | > 100 |

Notes: This table describes the distribution of $m_{ikt}^{jl} \equiv s_{ikt}^{[j]}/s_{ikt}^{[l]}$ where $j < l$. Only product × country pairs with positive and significant (at the 1 percent level) $\beta$s are considered.

Table 8: Comparative Advantage in Quality Production and Average Expenditure Shares

It is clearly seen that the $m_{ikt}^{jl}$'s mostly exceed one by a large margin. The median relative ranked expenditure share is between 23 and 90 depending on the cutoff for the number of importers per country-product used. Moreoever, the share of variety-firm pairs for which the high-$\beta$ variety accounts for a higher share of expenditure is generally small and does not exceed 11%. For example, when $\beta$ varies at the 4-digit level and we require at least 50 firms per product-country to estimate them, only 1.8% of the $m_{ikt}^{jl}$'s are below unity. Hence, on average, firms tend to allocate their spending towards low-$\beta$ countries. Together with the finding that large firms bias their spending towards these countries even more, shown in Table 6, this explains the systematic departure of the homotheticity property found in Section 3.

The fact that countries with a comparative advantage in high quality production are also large sellers overall suggests that the cross-country variation in quality elasticities $\beta$'s is partially systematic and correlated with the level of productivity $A_c$.[31] To provide evidence in this direction, we project the estimated $\beta$'s on country GDP per capita according to

$$\ln \beta_{ck} = \alpha_k + \phi \times \ln y_c + u_{ck}. \tag{31}$$

Here $\ln y_c$ denotes (log) GDP per capita in country $c$ and $\alpha_k$ is a product fixed effect so that the coefficient $\phi$ is identified from the variation across countries within narrowly-defined products. The resulting estimates are contained in Table 9. We see that, in most specifications, income per capita is indeed negatively related to the quality elasticity $\beta$. This suggests that rich countries have a comparative advantage in the production of high quality inputs.[32]

## 4.3 Taking Stock

In this section, we presented a theory of quality choice that endogenously generates a non-homothetic import demand system. We assumed that firm productivity and input quality are complements and that countries differ in their efficiency to produce high quality inputs. As predicted by this theory, we provided evidence that large firms pay higher prices for their inputs within countries, that this price-size elasticity varies substantially

[31] Note that, holding $\beta$ fixed, the expenditure share on country $c$ is increasing in $A_c$ for all firms - see Proposition 3.

[32] While reminiscent, this result differs from the findings in Khandelwal (2010). Khandelwal (2010) infers a country-level quality index from aggregate trade flows across-countries. We use within-country variation across firms to estimate the quality elasticity $\beta_{ck}$ and then show that such elasticities are systematically correlated with income per capita.

| Dep. Var.: | ln $\beta_{ck}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Aggregation level: | 8 digit products (fine) | | | | 4 digit products (coarse) | | | |
| Sample of betas: | Positive sig. | Positive sig. | Positive sig. | Positive | Positive sig. | Positive sig. | Positive sig. | Positive |
| Min obs./variety: | 10 | 50 | 100 | 50 | 10 | 50 | 100 | 50 |
| ln GDP/cap. | -0.173 | -0.160** | -0.201*** | 0.005 | -0.404*** | -0.114 | -0.137* | -0.149* |
| | (0.127) | (0.064) | (0.055) | (0.090) | (0.142) | (0.068) | (0.079) | (0.088) |
| Nb varieties (obs. 1st stage) | 2,809 | 755 | 429 | 755 | 1,701 | 1,004 | 801 | 1,004 |
| Nb products | 1,836 | 469 | 254 | 469 | 637 | 393 | 308 | 393 |
| Nb countries (obs. 2nd stage) | 52 | 28 | 19 | 47 | 67 | 43 | 35 | 67 |
| $R^2$ | 0.06 | 0.15 | 0.15 | 0.00 | 0.20 | 0.04 | 0.08 | 0.07 |

Notes: This tables reports the correlations obtained in a regressions of $\ln \beta_{ck}$ on product and country fixed effects, and (ln) country level GDP per capita. The subset of $\beta_{ck}$ coefficients that are significant at the 1 percent level in first stage regressions is denoted by "sig." All regressions are weighted by import values. Robust standard errors in parentheses with ***,** and * respectively denoting significance at the 1%, 5% and 10% levels.

Table 9: Comparative Advantage in Quality Production and GDP per capita

across countries, and that large firms concentrate their spending on countries where this elasticity is small. In this way, this theory of quality choice provides a mechanism to rationalize the finding, documented in Section 3, that large firms feature more concentrated import expenditure shares.

One can certainly think of other mechanisms to explain the findings of Section 3. One plausible example is costly supplier search, if larger firms tend to connect with better and more expensive input suppliers in every country. Similarly, high input prices could reflect the remuneration for relationship-specific investments that suppliers undertake if their customers are large. To generate the specific of pattern of non-homothetic import demand found in Section 3, however, these theories would need to explain why such mechanisms are more pronounced in firms' most important sourcing countries.

# 5 Conclusion

In this paper, we use data for French manufacturing firms and document that the demand for imported inputs is non-homothetic. In particular, holding the extensive margin of trade constant, large firms tend to concentrate their import spending on their top sourcing countries. This finding is inconsistent with virtually all firm-based models of importing that have been considered in the literature. These theories assume that productivity is factor neutral and hence does not affect the allocation of expenditure across a given set of countries.

To rationalize this feature of the data, we then propose a novel model of importing with input quality choice. In particular, we introduce two novel ingredients into an otherwise standard firm-based model of input trade. First, we assume that countries can produce their input in a variety of qualities. Second, we allow for a complementarity between firm-productivity and input-quality. This model gives rise to a non-homothetic import demand system, as long as countries differ in their comparative advantage in producing high quality inputs. More precisely, productive firms bias their import spending towards efficient quality producers. We provide evidence in support of this and other predictions of this model. A crucial step of our methodology is to estimate the cost of producing quality across countries using the within-country relationship between input prices and firm size.

At a theoretical level, we provide an equivalence result showing that this micro-founded model of quality

choice has an "Armington-type" representation, which extends the baseline model of importing in a single dimension: firm productivity is no longer factor neutral but biased towards particular countries. Whether or not this parsimonious departure from the model used in the literature is quantitatively important for macroeconomic questions, such as the gains from trade or explaining aggregate trade flows, is an interesting direction for future research.

# References


**Acemoglu, Daron**, "Equilibrium Bias of Technology," *Econometrica*, 2007, *75* (5), 1371–1409.

**Amiti, Mary, Oleg Itskhoki, and Jozef Konings**, "Importers, Exporters, and Exchange Rate Disconnect," *American Economic Review*, 2014, *104* (7), 1942–1978.

**Antràs, Pol, Teresa C. Fort, and Felix Tintelnot**, "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms," *American Economic Review*, 2017, *107* (9), 2514–64.

**Armenter, Roc and Miklós Koren**, "A Balls-and-Bins Model of Trade," *American Economic Review*, 2014, *104* (7), 2127–51.

**Blaum, Joaquín**, "Global Firm in Large Devaluations," *Unpublished Manuscript, Brown University*, 2019.

**Blaum, Joaquin, Claire Lelarge, and Michael Peters**, "The Gains from Input Trade with Heterogeneous Importers," *American Economic Journal: Macroeconomics*, 2018, *10* (4), 77–127.

**Correia, Sergio**, "Singletons, Cluster-Robust Standard Errors and Fixed Effects: A Bad Mix," Mimeo, Duke University 2015.

**Eaton, Jonathan, Samuel Kortum, and Francis Kramarz**, "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 2011, *79* (5), 1453–1498.

**Eslava, Marcela, Ana Cecilia Fieler, and Daniel Yi Xu**, "Trade, Skills, and Quality Upgrading: A Theory with Evidence from Colombia," *American Economic Review*, 2017. forthcoming.

**Fieler, Ana Cecilia**, "Non-Homotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, 2011, *79* (4), 1069–1101.

**Gaure, Simen**, "OLS with Multiple High Dimensional Category Variables," *Computational Statistics & Data Analysis*, 2013, *66*, 8 – 18.

**Goldberg, Pinelopi, Amit Khandelwal, Nina Pavcnik, and Petia Topalova**, "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *Quarterly Journal of Economics*, 2010, *125* (4), 1727–1767.

**Gopinath, Gita and Brent Neiman**, "Trade Adjustment and Productivity in Large Crises," *American Economic Review*, 2014, *104* (3), 793–831.

**Halpern, Laszlo, Miklos Koren, and Adam Szeidl**, "Imported Inputs and Productivity," *American Economic Review*, 2015, *105* (12), 3660–3703.

**Hillberry, Russell and David Hummels**, "Trade Elasticity Parameters for a Computable General Equilibrium Model," in "Handbook of Computable General Equilibrium Modeling," Vol. 1, Elsevier, 2013, pp. 1213–1269.

**Hsieh, Chang-Tai and Peter J. Klenow**, "Relative Prices and Relative Prosperity," *American Economic Review*, 2007, *97* (3), 562–585.

**Johnson, Robert C and Guillermo Noguera**, "A Portrait of Trade in Value-Added over Four Decades," *Review of Economics and Statistics*, 2017, *99* (5), 896–911.

**Kasahara, Hiroyuki and Joel Rodrigue**, "Does the Use of Imported Intermediates Increase Productivity? Plant-Level Evidence," *Journal of Development Economics*, 2008, *87*, 106–118.

**Khandelwal, Amit**, "The Long and Short (of) Quality Ladders," *The Review of Economic Studies*, 2010, *77* (4), 1450–1476.

**Kugler, Maurice and Eric Verhoogen**, "Plants and Imported Inputs: New Facts and an Interpretation," *American Economic Review (P&P)*, 2009, *99* (2), 501–507.

**_ and _** , "Prices, Plant Size, and Product Quality," *Review of Economic Studies*, 2012, *79* (1), 307–339.

**Lu, Dan, Asier Mariscal, and Luis-Fernando Mejia**, "How Firms Accumulate Inputs: Evidence from Import Switching," 2016. Working Paper.

**Manova, Kalina and Zhiwei Zhang**, "Export Prices Across Firms and Destinations," *Quarterly Journal of Economics*, 2012, *127*, 379–436.

**Miroudot, Sébastien, Rainer Lanz, and Alexandros Ragoussis**, "Trade in Intermediate Goods and Services," OECD Trade Policy Papers 93 2009.

**Petrin, Amil and James Levinsohn**, "Measuring Aggregate Productivity Growth Using Plant-Level Data," *RAND Journal of Economics*, 2012, *43* (4), 705–725.

**Ramanarayanan, Ananth**, "Imported Inputs and the Gains from Trade," 2014. Working Paper.

**Waugh, Michael E.**, "International Trade and Income Differences," *American Economic Review*, 2010, *100* (5), 2093–2124.

# APPENDIX

# A Theoretical Appendix

## A.1 Proof of Proposition 1

Consider the minimization problem in (6). Consider the problem in efficiency units, defined as $\tilde{z}_{ck} \equiv \eta_{ck} z_{ck}$. The first order conditions are given by

$$\frac{1}{\xi_{ck}} = \lambda\varphi \frac{\partial f(x)}{\partial x_k} \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{ck}}, \tag{32}$$

where $\lambda$ is the multiplier on the constraint. Consider a product $k$. Then

$$\sum_{c\in\Sigma_k} \tilde{z}_{ck}/\xi_{ck} = \lambda\varphi \frac{\partial f(x)}{\partial x_k} \sum_{c\in\Sigma_k} \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{ck}} \tilde{z}_{ck} = \lambda\varphi \frac{\partial f(x)}{\partial x_k} x_k,$$

where the last equality follows from constant returns of $\hat{g}_k$.[33] Summing across products,

$$\Gamma(\Sigma, x, \varphi) = \sum_{\{c\in\Sigma_k\}_k} \tilde{z}_{ck}/\xi_{ck} = \lambda\varphi \sum_{k=1}^{n} \frac{\partial f(x)}{\partial x_k} x_k = \lambda\varphi f(x) = \lambda x, \tag{33}$$

so that $\lambda$ is the marginal cost of production. The expenditure share of variety 1 in product $k$ is:

$$s_{1k} = \frac{\tilde{z}_{1k}/\xi_{1k}}{\sum_{c\in\Sigma_k} \tilde{z}_{ck}/\xi_{ck}} = \frac{\frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{1k}} \tilde{z}_{1k}}{\hat{g}_k(\tilde{z}_{1k}, \tilde{z}_{2k}, ..., \tilde{z}_{C_k k})} = \frac{\frac{\partial}{\partial \tilde{z}_{1k}} \hat{g}_k(1, \chi_{2k}, ..., \chi_{C_k k})}{\hat{g}_k(1, \chi_{2k}, ..., \chi_{C_k k})},$$

where $\chi_{ck} \equiv \tilde{z}_{ck}/\tilde{z}_{1k}$ is the relative quantity of country c and $C_k \equiv |\Sigma_k|$ is the number of countries in the sourcing strategy for product k. Thus, the expenditure share on any particular country is a function of relative quantities $[\chi_{ck}]$. We now show that $[\chi_{ck}]$ depend only on prices $[\xi_{ck}]$. (32) implies that

$$\left[\lambda\varphi \frac{\partial f(x)}{\partial x_k}\right]^{-1} = \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{c'k}} \xi_{c'k} = \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{ck}} \xi_{ck} \text{ for all } c, c' \in \Sigma_k. \tag{34}$$

As $\hat{g}_k$ is constant returns to scale, (34) implies that

$$\frac{\partial \hat{g}_k(1, \chi_{2k}, .., \chi_{C_k k})}{\partial \tilde{z}_{ck}} = \frac{\partial \hat{g}_k(1, \chi_{2k}, .., \chi_{C_k k})}{\partial \tilde{z}_{1k}} \frac{\xi_{1k}}{\xi_{ck}} \text{ for all } c \in \Sigma_k. \tag{35}$$

These are $|\Sigma_k|-1$ equation in $|\Sigma_k|-1$ unknowns $(\chi_{2k}, ..., \chi_{C_k k})$ which have a solution $\chi_j = \rho_j\left([\xi_{ck}]_{c\in\Sigma_k}\right)$ for all $j \neq 1$. Hence, we have that

$$s_{ck}(\Sigma, x, \varphi) = h_k\left([\xi_{ck}]_{c\in\Sigma_k}\right).$$

[33] To see this note that $\hat{g}_k(\mu\tilde{z}_{1k}, ..., \mu\tilde{z}_{nk}) = \mu\hat{g}_k(\tilde{z}_{1k}, ..., \tilde{z}_{nk})$. Differentiating with respect to $\mu$ yields

$$\sum_c \frac{\partial \hat{g}_k\left(\mu\tilde{z}_{1k}, ..., \mu\tilde{z}_{C_k k}\right)}{\partial(\mu\tilde{z}_{ck})} \tilde{z}_{ck} = \hat{g}_k\left(\tilde{z}_{1k}, ..., \tilde{z}_{C_k k}\right).$$

Evaluating this condition for $\mu = 1$ yields the result.

## A.2 Proof of Proposition 2

Consider the maximization problem in (17). Using the definition of $\eta$, the Lagrangian is given by

$$\mathcal{L} = \int_{q=0}^{\infty} \left(\gamma (\varphi)^{\theta} + (1-\gamma) (q)^{\theta}\right)^{1/\theta} z(q)\, dq + \lambda \left(X - \int_{q=0}^{\infty} p(q)\, z(q)\, dq\right),$$

where $X$ denotes the total amount of spending and the country $c$ subscript is suppressed for simplicity. The optimality conditions are given by

$$\left(\gamma (\varphi)^{\theta} + (1-\gamma) (q)^{\theta}\right)^{1/\theta} - \lambda p(q) \leq 0 \text{ for all } q$$

with equality whenever $z(q) > 0$. Given the linearity of the problem, there is generically only a single quality level that is bought. Call this quality level $q(\varphi)$. This quality level is defined as

$$\begin{aligned} q(\varphi) &= \arg\max_{q} \left\{ \frac{\left(\gamma (\varphi)^{\theta} + (1-\gamma) (q)^{\theta}\right)^{1/\theta}}{p(q)} \right\} = \arg\max_{q} \left\{ \frac{\left(\gamma (\varphi)^{\theta} + (1-\gamma) (q)^{\theta}\right)^{1/\theta}}{q^{\beta}} \right\} \\ &= \arg\max_{q} \left\{ \left(\gamma (\varphi)^{\theta} q^{-\beta\theta} + (1-\gamma) q^{(1-\beta)\theta}\right)^{1/\theta} \right\}. \end{aligned} \tag{36}$$

Note that $\gamma (\varphi)^{\theta} q^{-\beta\theta} + (1-\gamma) q^{(1-\beta)\theta}$ is decreasing in $q$ if $\beta > 1$. Below we show that $q(\varphi)$ is unique and given by

$$q(\varphi) = \left(\frac{\beta}{1-\beta}\frac{\gamma}{1-\gamma}\right)^{1/\theta} \varphi, \tag{37}$$

if and only if $\theta < 0$. Hence, $q(\varphi)$ is increasing in $\varphi$. Moreover, we have that

$$p(\varphi) = \frac{1}{A} q(\varphi)^{\beta} = \frac{1}{A} \left(\frac{\beta}{1-\beta}\frac{\gamma}{1-\gamma}\right)^{\frac{\beta}{\theta}} (\varphi)^{\beta}. \tag{38}$$

The aggregate import flow of spending $X$ units on the optimal quality vintage $q(\varphi)$, is given by

$$\left(\gamma (\varphi)^{\theta} + (1-\gamma) q(\varphi)^{\theta}\right)^{1/\theta} z(q(\varphi)) = \left(\gamma (\varphi)^{\theta} + (1-\gamma) q(\varphi)^{\theta}\right)^{1/\theta} \frac{X}{p(q(\varphi))}.$$

Substituting (37) and (38) we get that

$$\begin{aligned} \frac{\left(\gamma (\varphi)^{\theta} + (1-\gamma) q(\varphi)^{\theta}\right)^{1/\theta}}{p(q(\varphi))} X &= \frac{\left(\gamma (\varphi)^{\theta} + (1-\gamma) \frac{\beta}{1-\beta}\frac{\gamma}{1-\gamma} (\varphi)^{\theta}\right)^{1/\theta}}{\frac{1}{A}\left(\frac{\beta}{1-\beta}\frac{\gamma}{1-\gamma}\right)^{\beta/\theta} \varphi^{\beta}} X \\ &= \gamma^{1/\theta} \frac{\left(\frac{1}{1-\beta}\right)^{1/\theta}}{\left(\frac{\beta}{1-\beta}\frac{\gamma}{1-\gamma}\right)^{\beta/\theta}} A \varphi^{1-\beta} X \equiv \eta_c(\varphi) \times X. \end{aligned}$$

Note that (38) implies a log-linear relationship between prices and firm productivity, i.e.

$$\ln p_c(\varphi) = \beta_c \times \ln \varphi + \ln\left(\left(\frac{\beta_c}{1-\beta_c}\frac{\gamma}{1-\gamma}\right)^{\beta_c/\theta}\frac{1}{A_c}\right) = \beta_c \times \ln \varphi + \delta_c. \tag{39}$$

**Uniqueness of $q(\varphi)$** The optimal quality choice is given in as $q(\varphi) = \arg\max_q G(q,\varphi)$ where

$$G(q,\varphi) = \left(\gamma\varphi^\theta q^{-\beta\theta} + (1-\gamma)\, q^{(1-\beta)\theta}\right)^{1/\theta}$$

We now show that $q(\varphi)$ is unique if $\theta < 0$. Note that

$$\begin{aligned}\frac{\partial G(q,\varphi)}{\partial q} &= \frac{1}{\theta}G(q,\varphi)^{1-\theta}\left((-\beta\theta)\gamma\varphi^\theta q^{-\beta\theta-1} + (1-\beta)\theta(1-\gamma)\, q^{(1-\beta)\theta-1}\right)\\ &= G(q,\varphi)^{1-\theta}\left(-\beta\gamma\varphi^\theta q^{-\beta\theta-1} + (1-\beta)(1-\gamma)\, q^{(1-\beta)\theta-1}\right).\end{aligned}$$

Hence, the optimal $q(\varphi)$ solves

$$\beta\gamma\varphi^\theta q^{-\beta\theta-1} = (1-\beta)(1-\gamma)\, q^{(1-\beta)\theta-1}. \tag{40}$$

The second derivative can be written as

$$\begin{aligned}\frac{\partial^2 G(q,\varphi)}{\partial q^2} &= (1-\theta)G(q,\varphi)^{-\theta}\frac{\partial G(q,\varphi)}{\partial q}\left(-\beta\gamma\varphi^\theta q^{-\beta\theta-1} + (1-\beta)(1-\gamma)\, q^{(1-\beta)\theta-1}\right)\\ &\quad + G(q,\varphi)^{1-\theta}\left(\beta(\beta\theta+1)\gamma\varphi^\theta q^{-\beta\theta-2} + ((1-\beta)\theta-1)(1-\beta)(1-\gamma)\, q^{(1-\beta)\theta-2}\right).\end{aligned}$$

Evaluating this expression at $q(\varphi)$ yields

$$\begin{aligned}\frac{\partial^2 G(q(\varphi),\varphi)}{\partial q^2} &= G(q,\varphi)^{1-\theta} q^{-1}\left(\beta(\beta\theta+1)\gamma\varphi^\theta q^{-\beta\theta-1} + ((1-\beta)\theta-1)(1-\beta)(1-\gamma)\, q^{(1-\beta)\theta-1}\right)\\ &= G(q,\varphi)^{1-\theta}\theta\gamma\varphi^\theta\beta q^{-\beta\theta-2}.\end{aligned}$$

Hence, $\frac{\partial^2 G(q(\varphi),\varphi)}{\partial q^2} < 0$ if and only if $\theta < 0$. Note also that $G(q(\varphi),\varphi) = \left(\frac{1-\gamma}{\beta}\right)^{1/\theta} q^{1-\beta} > 0$ and that $\lim_{q\to 0} G(q,\varphi) = \lim_{q\to\infty} G(q,\varphi) = 0$. Hence, (40) implicitly defines the optimal quality choice as long as $\theta < 0$. Rearranging (40) yields

$$q(\varphi) = \left(\frac{\beta}{1-\beta}\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\theta}} \times \varphi.$$

## A.3 Proof of Proposition 3

We show that the two economies generate exactly the same import demand system across foreign varieties. Let $M$ be the level of spending on imports. The import demand system from the economy described in Proposition 3 is defined by

$$z(\varphi) \equiv \arg\min_{[z_c]_{c=1}^{\Sigma_i}}\left\{\left(\sum_{c=1}^{\Sigma_i}\left(\varphi^{1-\beta_c} z_c\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}} \text{ s.t. } \sum_{c=1}^{\Sigma_i}\frac{1}{Z_c}z_c \le M\right\}. \tag{41}$$

Now consider the model with input quality choice. The input demand across varieties is defined by

$$\min_{[z_c(q)]_{c=1}^{\Sigma_i}} \left\{ \left( \sum_{c=1}^{\Sigma_i} \left( \int_{q\in\mathcal{Q}_c} \eta\left(q,\varphi\right) z_c\left(q\right) dq \right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \text{ s.t. } \sum_{c=1}^{\Sigma_i} \int_{q\in\mathcal{Q}_c} p\left(q,\varphi\right) z_c\left(q\right) dq \leq M \right\} \tag{42}$$

Substituting the optimal choices characterized in proposition 2, we get that

$$\min_{[z_c]_{c=1}^{\Sigma_i}} \left\{ \left( \sum_{c=1}^{\Sigma_i} \left( \eta_c\left(\varphi\right) z_c \right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \text{ s.t. } \sum_{c=1}^{\Sigma_i} p_c\left(\varphi\right) z_c \leq M \right\}, \tag{43}$$

where $p_c\left(\varphi\right) = \frac{1}{A_c}\left(\frac{\beta_c}{1-\beta_c}\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\theta}\beta_c} \varphi^{\beta_c}$ and $\eta_c\left(\varphi\right) = \left(\frac{\gamma}{1-\beta_c}\right)^{1/\theta} \varphi$. Hence,

$$\min_{[z_c]_{c=1}^{\Sigma_i}} \left\{ \left( \sum_{c=1}^{\Sigma_i} \left( \left( \frac{\gamma}{1-\beta_c} \right)^{1/\theta} \varphi z_c \right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \text{ s.t. } \sum_{c=1}^{\Sigma_i} \frac{1}{A_c} \left( \frac{\beta_c}{1-\beta_c}\frac{\gamma}{1-\gamma} \right)^{\frac{1}{\theta}\beta_c} \varphi^{\beta_c} z_c \leq M \right\} \tag{44}$$

$$= \min_{[\tilde{z}_c]_{c=1}^{\Sigma_i}} \left\{ \left( \sum_{c=1}^{\Sigma_i} \left( \varphi^{1-\beta_c} \tilde{z}_c \right)^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} \text{ s.t. } \sum_{c=1}^{\Sigma_i} \frac{1}{A_c} \left[ \left( \frac{\beta_c}{1-\gamma} \right)^{\beta_c} \left( \frac{1-\beta_c}{\gamma} \right)^{1-\beta_c} \right]^{1/\theta} \tilde{z}_c \leq M \right\}, \tag{45}$$

where $\tilde{z}_c = \left(\frac{\gamma}{1-\beta_c}\right)^{1/\theta} z_c \varphi^{\beta_c}$. This is the same problem as (41) if $Z_c$ is defined as in (23).

# B Empirical Appendix

## B.1 Data Description

Our main data set stems from the information system of the French custom administration (DGDDI) and contains the majority of import and export flows by French firms, although we focus in the present paper on French manufacturing firms. The data is collected at the 8-digit (NC8) level. A firm located within the French metropolitan territory must report detailed information as long as the following criteria are met. For imports from outside the EU, reporting is required from each firm and flow if the imported value exceeds 1,000 Euros. For within EU imports, import flows have to be reported as long as the firm's annual trade value exceeds 100,000 Euros.[34] However, some firms that are below the threshold (ca. 80,000 firm-year observations out of ca. 150,000 reporting intra-EU imports) voluntarily report, most probably because they are simultaneously subject to detailed reporting of their exports to EU countries.

In spite of this partial censoring of the data, the attractive feature of the French data is the presence of unique firm identifiers (the SIREN code) that is available in all French administrative files. Hence, various datasets can be matched to the trade data at the firm level. The fiscal files provide rich complementary accounting information. They consist of two different files: the BRN ("Bénéfices Réels Normaux") and the RSI ("Régime Simplifié d'Imposition"). The BRN contains the balance sheet of all firms in the traded sectors with sales above 730,000 Euros. The RSI is the counterpart of the BRN for firms with sales below 730,000 Euros. Although the details of the reporting differs, for our purposes these two data sets contain

[34] This threshold was in effect between 2001 and 2006, which is period we focus on. Between 1993 and 2001, the threshold was ca. 40,000 Euros. After 2006, it was raised to 150,000 Euros and to 460,000 Euros after 2011.

essentially the same information. Their union covers nearly the entire universe of French firms. To measure the expenditure on domestic inputs, we simply subtract the total import value from the total expenditure on wares and inputs reported in the fiscal files. Capital is measured at book value (historical cost).

Finally, we incorporate information on the ownership structure from the LIFI/DIANE (BvDEP) files. These files are constructed at INSEE using a yearly survey (LIFI) that describes the structure of ownership of all firms in the private sector whose financial stakes in other firms are higher than 1.2 million euros or whose sales are above 60 million euros or with employment higher than 500 employees. This survey is complemented with the information about ownership structure available in the DIANE (BvDEP) files, which are constructed using the annual mandatory reports to commercial courts and the register of firms that are controlled by the State.

Using these three sets of data, we construct a non-balanced panel dataset spanning the period from 2001 to 2006. Some basic descriptive statistics are contained in Table 1 in the main text. Our sample includes a total of 187,191 firm $\times$ year observations, which implies that there are about 31,000 manufacturing importers per year. The median importing firm has 23 workers and annual sales of slightly more than 3 millions euros. Three third of importing firms are also exporters, and this proportion raises to 86% for "large" importers, i.e. firms importing several varieties of a least one product.

## B.2 Additional Empirical Results and Robustness

In this section we report various extensions and robustness checks for the baseline empirical results that are presented in the main text.

### B.2.1 The Variety Margin

**Import Trade Interactions.** Table 10 complements Table 1 and contains descriptive statistics on trade flows at the variety level.[35] Table 10 shows that our custom data are sliced into (a bit less than) 800,000 variety-firm pairs per year. Given that there are about 31,000 importers per year in our data (Table 1), this implies that the average importer imports about 25 varieties of potentially different products. This average masks, however, substantial heterogeneity. The median country is only active in 68 firm-product cells, whereas the top two exporting countries to France, namely Germany and Italy, report ca. 160,000 and 100,000 interactions per year respectively in distinct firm $\times$ product cells. Similarly, for half of the potential products (i.e. roughly 5,000 products) only 28 country $\times$ firm interactions are observed, while the most popular products are shipped into France in more than 800 distinct country-firm combinations.[36]

Finally, the two remaining rows confirm the findings of Gopinath and Neiman (2014) for the case of Argentina that imports are also very concentrated at the firm level. While the median firm sources only 8 varieties a year internationally, the top one percent of firms (ca. 300 firms) import 246 varieties. Similarly, while the most "valuable" firm $\times$ variety pairs are worth more than several millions Euros, a quarter of French importers import less than 1,000 Euros worth of the varieties within a year.

**Distribution of Spending.** How do firms allocate their expenditure (for one considered product) across source countries, i.e. across varieties? Figure 5 complements the two last rows of Table 10 and depicts the

[35] The statistics in Table 10 are reminiscent of the discussion of sparsity in Armenter and Koren (2014). We note, however, that they analyze the data at the *flow* level, while we aggregate it to the *firm-variety* level. It is this dimension our model can speak to.

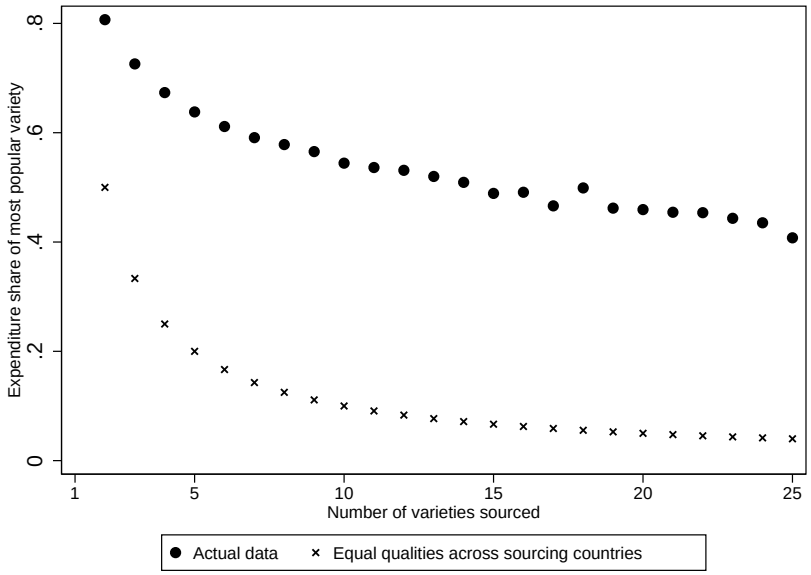
[36] Note in particular the still very large difference between the $99^{\text{th}}$ and $90^{\text{th}}$ quantiles.

| Variable | Dimension | Quantiles 25th | 50th | 75th | 90th | 99th |
|---|---|---|---|---|---|---|
| (1) # of annual trade interactions per country | firms × products | 14 | 68 | 478 | 4,722 | 74,510 |
| (2) # of annual trade interactions per product | firms × countries | 9 | 28 | 79 | 191 | 839 |
| (3) # of annual trade interactions per importing firm | products × countries | 2 | 8 | 25 | 60 | 246 |
| (4) Value of firm x product x country annual shipments | thousand euros | 1 | 6 | 35 | 168 | 2,765 |
| Average number of annual firm × product × country observations: | | | | | | 790,837 |
| Total number of (year × firm × product × country) observations: | | | | | | 4,745,021 |

Notes: French importing firms active in the manufacturing industries, between 2001 and 2006. Values of shipments are expressed in thousand euros. See Section B.1 in the Appendix for a complete description of the data.

Table 10: The Concentration of French Imports at the Variety Level

average expenditure share on the main variety for firms sourcing $V$ varieties per product, for different values of $V$ on the $x$-axis. This figure essentially shows that the intensive margin of importing exhibits substantial concentration since (in particular) most importers concentrate their spending heavily on their most important variety.




Notes: The figure shows the average expenditure share on the top variety (y-axis) of firm×product pairs with varying number of varieties (x-axis). We also depict the counterfactual expenditure share if price-adjusted quality flows (and hence expenditure shares) were equalized across varieties. We use 6 years of data from 2001-2006 and report the yearly average.

Figure 5: Concentration of Firms' Import Spending on Main Variety Across Sourcing Strategies

Firms sourcing as many as 10 varieties spend on average more than 50 percent of their total expenditure on their main variety. Similarly, firms with 25 varieties per product still spend 40% of their import spending on their top trading partner. For comparison, we also depict the expenditure share if spending was equalized across varieties. It is clearly seen that this expenditure share is counterfactually low. Thus, importers rely heavily on a small set of "core" suppliers for most of their import budget, while a multitude of marginal sourcing countries seem to play a more minor role.

### B.2.2 Disagreement on the Intensive Margin

We formalize this notion of disagreement through a variance decomposition exercise. We measure how much of the variation in firms' expenditure share in their top supplying country is accounted by the extensive margin of trade. Formally, we estimate

$$\ln\left(\frac{s_{ikt}^{(1)}}{1-s_{ikt}^{(1)}}\right) = \alpha_S + \alpha_t + \alpha_k + \alpha_{\Sigma_{ki}} + u_{ikt}, \tag{46}$$

where $s_{ikt}^{(1)}$ is firm $i$'s expenditure share on its top variety within 8-digit product $k$ and $\alpha_S$, $\alpha_t$ and $\alpha_k$ are sector, time and product fixed effects. Importantly, $\alpha_{\Sigma_{ki}}$ denotes a product-specific control for firms' sourcing strategy. We control for the sourcing strategy via the interaction of product fixed effects and the set of supplying countries. We drop any sourcing strategy that has a single observation because theory cannot be tested on these observations.

Table 11 contains the results of estimating (46). We find ample variation in firms' expenditure shares after controlling for the sourcing strategy. The entirety of about 97,000 fixed effects explain 43.5% of the variation in the data (excluding singletons).[37] When we allow the coefficients to depend on the size of the sourcing set and - in the same spirit as in the right panel of Figure 3 - distinguish different samples by the cardinality of sourcing strategies, the explanatory power increases but never exceeds 70% of the variation across non-singleton observations. Overall, the findings in Figure 3 and Table 11 indicate that firms' sourcing strategies are not a sufficient statistic for the distribution of foreign spending as predicted by standard models of importing.

| Dependent variable: | $\ln\left(\frac{s^{(1)}}{1-s^{(1)}}\right)$ | | | | |
|---|---|---|---|---|---|
| Nb. varieties: | $\geq 2$ | 2 | 3 | 4 | 5 |
| | (1) | (2) | (3) | (4) | (5) |
| Observations | 739,167 | 432,594 | 147,921 | 65,930 | 34,355 |
| Identified FE | 403,073 | 165,76 | 98,004 | 53,566 | 30,678 |
| of which singletons | 306,120 | 101,549 | 76,498 | 46,191 | 27,904 |
| Variance decomposition, removing singletons (Correia, 2015; Gaure, 2013): | | | | | |
| $R^2$ | 0.435 | 0.362 | 0.493 | 0.602 | 0.707 |
| Controls, year and ind. FE | 0.014 | 0.014 | 0.033 | 0.125 | 0.410 |
| Product × Sourcing Strat. FE | 0.425 | 0.351 | 0.486 | 0.648 | 0.956 |
| Covariance term | -0.005 | -0.003 | -0.026 | -0.172 | -0.660 |

Notes: Regressions at the importer × product level, importers with more than two, exactly two, three, four or five varieties depending on the specification, observed from 2001 to 2006. All regressions include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating that the group has foreign headquarters, and an indicator of capital intensity (tangible assets per worker). All regressions include fixed effects for years and four-digit industries. Observations are weighted to give each firm an equal weight.

Table 11: Within Sourcing-Strategy Variance of Firms' Expenditure Shares

[37] Singletons are observations whose sourcing strategy is unique. To be precise, the sample without singletons in column 1 of Table 11 is comprised of 739,167 - 306,120 = 433,047 observations. 403,073 - 306,120 = 96,953 fixed effects actually control for product × sourcing strategies in this subsample.

### B.2.3 The Intensive Margin of Imports: Controlling for the Number of Trading Partners

In our main analysis in Tables 2 and 3 in Section 3 we control for firms' sourcing strategies in an unrestricted way. In Tables 12 and 13 below, we replicate these results when we only control only for the number of sourcing countries. Doing so leaves all our results qualitatively and quantitatively unchanged.

| | ln Herf. | ln Gini | Expenditure share on... varieties | | | | |
|---|---|---|---|---|---|---|---|
| | | | Top 1 | Top 2 | Top 3 | Top 4 | Top 5 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| ln Sales | 0.018*** | 0.035*** | 0.010*** | 0.007*** | 0.006*** | 0.005*** | 0.005*** |
| | (0.001) | (0.002) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| ln Nb countries | -0.476*** | 0.813*** | -0.202*** | -0.181*** | -0.149*** | -0.125*** | -0.107*** |
| | (0.002) | (0.005) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| Impact of $\Delta_{sd}$ sales: | 0.047 | 0.093 | 0.026 | 0.018 | 0.013 | 0.012 | 0.010 |
| Sample mean: | -0.449 | -1.240 | 0.756 | 0.892 | 0.937 | 0.959 | 0.970 |
| Observations | 739,167 | 737,668 | 739,167 | 306,573 | 158,652 | 92,722 | 58,367 |
| Identified FE | 119,862 | 119,452 | 119,862 | 55,686 | 31,303 | 19,660 | 13,103 |
| of which singletons | 46,765 | 46,542 | 46,765 | 21,483 | 11,881 | 7,431 | 4,892 |

Notes: Regressions are estimated at the importer × product level. The estimation samples include importers with more than two, three, four, five or six varieties depending on the specification observed from 2001 to 2006. Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include fixed effects for years and 4-digit industries interacted with 8-digit products. All regressions also include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating foreign headquarters, and a measure of capital intensity (log tangible assets per worker). Observations are weighted to give each firm an equal weight independently of the number of imported products. $\Delta_{sd}$ sales denotes the marginal impact of an increase in sales by one standard deviation on the respective dependent variables. "Sample mean" refers to the sample mean of dependent variable.

Table 12: Non-homothetic Import Demand: Controlling for the Number of Trading Partners

| Aggregation level: | 8-dig. products | | | | 4-dig. products | |
|---|---|---|---|---|---|---|
| | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(4)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(5)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ln Sales | 0.107*** | 0.148*** | 0.174*** | 0.185*** | 0.102*** | 0.147*** |
| | (0.006) | (0.012) | (0.018) | (0.026) | (0.005) | (0.009) |
| ln Nb countries | -0.879*** | -1.540*** | -2.045*** | -2.493*** | -0.917*** | -1.658*** |
| | (0.015) | (0.030) | (0.051) | (0.079) | (0.012) | (0.023) |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | | | |
| - On dependent variable | 0.286 | 0.376 | 0.423 | 0.432 | 0.274 | 0.376 |
| - On $s^{(k)}$ | -0.041 | -0.022 | -0.013 | -0.008 | -0.038 | -0.022 |
| Sample average of $s^{(k)}$ | 0.206 | 0.080 | 0.041 | 0.024 | 0.204 | 0.081 |
| Observations | 739,167 | 306,573 | 158,652 | 92,722 | 597,628 | 283,009 |
| Identified FE | 462,159 | 226,593 | 126,607 | 77,729 | 315,037 | 180,785 |
| of which singletons | 339,672 | 181,470 | 105,941 | 67,223 | 211,584 | 133,194 |

Notes: Regressions are estimated at the importer × product level. The estimation samples include importers with more than two, three, four or five varieties depending on the specification, observed from 2001 to 2006. Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. The Regressions include fixed effects for years, 4-digit industries interacted with 8 or 4-digit products and sourcing countries (of the specific varieties considered). All regressions include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). Observations are weighted to give each firm an equal weight, independently of the number of imported products. The marginal impacts of an increase of sales by one standard deviation are computed at the mean of each sample in terms of sales, $s^{(1)}$, $s^{(2)}$, $s^{(3)}$, $s^{(4)}$ and $s^{(5)}$.

Table 13: Non-homothetic Import Demand: Controlling for the Number of Trading Partners in the CES Specification

### B.2.4 The Intensive Margin of Imports: Robustness and Extensions

In this section we report the results from a variety of robustness checks for our results in Section 3. In particular, we first show that our results reported in Table 3 are robust with respect to (i) the inclusion of firm fixed effects, (ii) different weighting schemes, (iii) different measures of firm size (e.g. employment or TFP), (iv) controlling for intra-firm trade, (v) different levels of aggregation both in the product and the time dimension and (vi) detailed geographic controls for firms' location within France to control for variation in market access across firms. To economize on space, Table 14, which contains the results for all these robustness tests, only reports the main coefficient of interest for all the different specifications. The detailed results are contained in Section O.2 of the Online Appendix.

**Firm Fixed Effects.** We first replicate the baseline analysis of Table 3 while incorporating firm level fixed effects into the estimated specifications. This implies that the coefficient of interest associated with sales is now identified from time *variation* in both firm size and import expenditure shares. The benefit of such specification is that they enable us to control for potential firm-level, time invariant unobserved heterogeneity, which is potentially correlated with both firm size and importing behavior.[38] Reassuringly, the obtained results are very similar to those of the baseline specification, in spite of the different identification strategy. The detailed results are contained in Table 18 in Section O.2 of the Online Appendix.

**Un-Weighted Regressions.** All of the regressions that are presented in the main text are weighted to give each firm an equal weight, independently of the number of imported products. The reason for this weighting procedure is that we do not want to over-emphasize the largest firms, which account for a disproportionate share of import flows, as documented in Table 10 and Figure 5. In Table 14 however, we present the results obtained for our central specification (Table 3) without weighting. The obtained results are almost unaffected, both qualitatively and quantitatively. The detailed results are contained in Table 19 in Section O.2 of the Online Appendix.

**Alternative Proxies for Productivity.** In our central specification of Table 3, firm sales are first and foremost a proxy for overall firm size, a firm level characteristic which enable to test Proposition 1. We now investigate whether our baseline results are robust to other proxies for firm productivity. In Table 14 we use first employment as an alternative measure of firm size.[39] We also present the results obtained with (revenue) productivity. We estimate revenue productivity from a translog production function in value added, with the method proposed in Petrin and Levinsohn (2012). The regression results show that our findings are largely preserved qualitatively and quantitatively. The detailed results are contained in Tables 20 and 21 in Section O.2 of the Online Appendix.

**Controlling for Intra-Firm Trade** One interpretation of our results could be that they reflect intra-firm trade. If large firms are more likely to be linked with their most important suppliers and buy larger quantities in such intra-firm transactions, one would indeed expect that large firms have more concentrated expenditure

[38] For example, larger firms might be located closer to large infrastructure such as harbors for reasons unrelated to the mechanisms of interest in this paper. This more favorable location (which is unobserved in our data) might simultaneously affect firms' importing strategies, a phenomenon we want to abstract from.

[39] In this regression, we control for firm level heterogeneity in technology by controlling for firms' material intensity, i.e. (ln) intermediate inputs per worker.

| | Dependent variable | | | |
|---|---|---|---|---|
| | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(4)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(5)}}\right)$ |
| (A) Controlling for firm fixed effects | | | | |
| ln Sales | 0.109*** | 0.110** | 0.313*** | 0.085 |
| | (0.019) | (0.048) | (0.071) | (0.122) |
| (B) Unweighted regressions | | | | |
| ln Sales | 0.125*** | 0.174*** | 0.185*** | 0.165*** |
| | (0.005) | (0.014) | (0.026) | (0.051) |
| (C) Employment as the measure of firm size | | | | |
| ln Employment | 0.077*** | 0.120*** | 0.113*** | 0.094 |
| | (0.008) | (0.019) | (0.038) | (0.063) |
| (D) TFP as the measure of firm size | | | | |
| ln TFP | 0.234*** | 0.285*** | 0.588*** | 0.583* |
| | (0.038) | (0.099) | (0.173) | (0.336) |
| (E) Controlling for intra-firm trade | | | | |
| ln Sales | 0.119*** | 0.178*** | 0.202*** | 0.177*** |
| | (0.008) | (0.019) | (0.039) | (0.064) |
| (F) Excluding firms in international group | | | | |
| ln Sales | 0.136*** | 0.190*** | 0.292*** | 0.296*** |
| | (0.011) | (0.029) | (0.054) | (0.103) |
| (G) Aggregation of products at 4-digit level | | | | |
| ln Sales | 0.115*** | 0.166*** | 0.172*** | 0.220*** |
| | (0.007) | (0.015) | (0.030) | (0.049) |
| (H) Aggregation across years (2001 - 2006) | | | | |
| ln Sales | 0.086*** | 0.155*** | 0.184** | 0.509*** |
| | (0.019) | (0.054) | (0.078) | (0.173) |
| (I) Controlling for firms' location in France | | | | |
| ln Sales | 0.110*** | 0.082* | 0.278*** | 0.068 |
| | (0.020) | (0.046) | (0.067) | (0.110) |
| (J) Removing neighboring sourcing countries | | | | |
| ln Sales | 0.145*** | 0.165*** | 0.254** | 0.176 |
| | (0.018) | (0.044) | (0.107) | (0.235) |

Notes: This table summarizes a set of robustness checks for our main result. For all specification we control for the sourcing strategy with sourcing strategy specific fixed effects. We refer to the main text for the details of the different specifications. The full results including the number of observations and the marginal effects are contained in the Online Appendix.

Table 14: Firm Size and the Intensive Margin of Trade: Robustness

shares. While this is a plausible mechanism, we now present additional results that our main results do not seem to be driven by such considerations.

Recall that all our baseline regressions already explicitly control for whether or not the importing firm is part of a international group, i.e. whether the firm has foreign affiliates. Our data, however, contains more detailed information. In particular, we observe the precise country in which the affiliate is located. In Panel (E) of Table 14 we therefore explicitly control for the set of countries the importing firm has an affiliate in. Doing so does not affect the results. In particular, we still find a positive relationship between firm size and the concentration of expenditure shares. As an alternative to controlling for the set of foreign affiliates we can also focus on the sample of firms, which do not have any foreign affiliates. Specifically, we replicate our analysis for the set of firms, which are not part of an international group. We find that our results are almost identical to our baseline results, i.e. even among firms without foreign affiliates import demand is non-homothetic in that firm size is positively correlated with the concentration of expenditure shares. The detailed results are contained in Tables 22 and 23 in Section O.2 of the Online Appendix.

**Aggregation Across Products and Years** In our main analysis we measure import flows annually at the 8-digit level. In this section we show that our results are robust to these choices. Consider first the classification of products. We replicate our main results when we measure products at the 4-digit level.[40] It is clearly seen that the results are qualitatively and quantitatively very similar. We then turn to the issue of time aggregation. If import flows are lumpy, we might see a large extent of time-variation in firms' import flows. While it is less clear why such pattern would systematically tend to generate a positive correlation between firm size and the concentration of expenditure shares, we nevertheless redid our analysis on "long-run" import flows, i.e. import flows, which are aggregated between 2001 and 2006. The results show the exact same pattern as the annual data: the concentration of expenditure share is increasing in firm size. The detailed results are contained in Tables 24 and 25 in Section O.2 of the Online Appendix.

**Internal Geography in France** Our baseline analysis does not control for the internal geography in France. If trade costs depend on distance and firms of different size are differentially allocated across space, our results might be driven by variation in firms' distances to their suppliers. In this section we address this concern. Our data does contain detailed information on the location of each firm within France. In particular we observe in which region, department and municipality the firm is located in. There are 24 regions, 102 department and 10,149 municipalities. We can exploit this information to control for firms' distance to their sourcing countries. In Table 14 we replicate our main results when we *interact* firms' sourcing strategies with their location measured at the municipality level. Hence, the coefficient on sales is only identified from firms in the same location who buy their inputs from the same sourcing countries. We also provide an additional robustness checks to address this concern. In particular, we focus on trading partners, where the variation in distance *within* France is arguable less important, namely trading partners, which are far away. To do so, we drop all import flows, with close-by countries (Spain, Italy, Germany, Luxembourg, Belgium and the UK). On this restricted sample we find the same non-homothetic pattern as for our baseline. The detailed results are contained in Tables 26 and 27 in Section O.2 of the Online Appendix.

[40] We also replicated the results at the 6-digit level. These regressions are available upon request.

### B.2.5 Regressions by Sourcing Strategy

In the specifications of Table 3, the coefficients on firm sales $\beta$ were identified by pooling firms with different sourcing strategies and controlling for sourcing strategy fixed effects. We now consider more demanding empirical specifications which restrict the sample to firms that perfectly agree on the sourcing strategy by product. More precisely, for each product, we consider firms that source only from the most popular set of countries of cardinality $V$. We then estimate the specification (12) pooling all firm $\times$ products of cardinality $V$, for $V = 2, .., 5$. Because we always control for product fixed effects, the coefficient of interest is identified from the variation across firms who share the exact same sourcing strategy.[41]

The results, shown in Table 15, are very similar to the ones of Table 3, both qualitatively and quantitatively. Most coefficients are significantly positive and their magnitude tends to increase with the order of the statistic, which implies that the distribution of expenditure across ranked sourcing countries of large firms first-order stochastically dominates the one of their smaller sized counterparts.

**The CES Approach (for Regressions by Sourcing Strategies).** One challenge of the statistical analysis of the intensive margin of imports in Section 3.3 is to control as accurately as possible for the extensive margin. In the baseline regressions of Table 3, we insert high-dimensional sets of fixed effects, while the approach underlying Table 15 focuses even more precisely on firms' intensive margin problem by considering only firms $\times$ products with strictly identical sourcing strategies. A major limitation of these approaches is, however, that the number of observations drops substantially, thus decreasing the statistical power of the test. This issue becomes especially pressing for the case of 4 or 5 varieties. To address this concern, we exploit in Table 16 a property of the CES demand system that allows us to increase the sample size while still controlling for the identity of firms' sourcing countries. More specifically, when the production function is CES, as in (5), the model of Section 2 implies that the log difference between expenditure shares of any two varieties $c$ and $c'$ (of any two order statistics $j$ and $j'$) is given by:

$$\ln\left(\frac{s_{ijk}}{s_{ij'k}}\right) = \ln\left(s_{ijk}\right) - \ln\left(s_{ij'k}\right) = \left(\rho - 1\right)\left(\ln\left(\xi_{jk}\right) - \ln\left(\xi_{j'k}\right)\right), \tag{47}$$

which is not only independent of any firm characteristic conditional on the sourcing strategy, but is even independent of the sourcing strategy itself, as the log-linear structure of expenditure shares in the CES case allows us to "difference out" the sourcing strategy.[42] While equation (47) requires stronger assumptions than the general expenditure share equation (7) of Section 2, it is very useful. Indeed, the key advantage of (47) over (12) is that it can be tested by pooling firms that source both $c$ and $c'$ but that may otherwise disagree in their sourcing strategy. This approach therefore results in an increase in sample size.

[41] Note that both the identity of countries can vary across products and that firms do not have to have the same sourcing strategy across products.

[42] Of course (47) is only defined for pairs of varieties which are actually sourced. That is, expression (47) is valid for $c, c' \in \Sigma_{ki}$, and not defined otherwise. This means that when testing for the exclusion restriction implicit in (47) for two particular varieties we need to restrict ourselves to the set of firms that source those two varieties. An advantage of the differencing approach, coupled with the CES assumption, is that we can pool together firms that differ in their sourcing strategy as long as they agree in sourcing a particular pair of varieties.

| | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(4)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(5)}}\right)$ |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Most popular 2-variety sourcing strategies | | | |
| ln Sales | 0.091*** | | | |
| | (0.007) | | | |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.225 | | | |
| - On $s^{(k)}$ | -0.033 | | | |
| Sample average of $s^{(k)}$ | 0.199 | | | |
| Observations | 83.778 | | | |
| Identified FE | 10.547 | | | |
| of which singletons | 2.189 | | | |
| | Most popular 3-variety sourcing strategies | | | |
| ln Sales | 0.058*** | 0.177*** | | |
| | (0.013) | (0.018) | | |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.135 | 0.414 | | |
| - On $s^{(k)}$ | -0.019 | -0.022 | | |
| Sample average of $s^{(k)}$ | 0.218 | 0.070 | | |
| Observations | 24.747 | 24.747 | | |
| Identified FE | 8,740 | 8,740 | | |
| of which singletons | 3.767 | 3.767 | | |
| | Most popular 4-variety sourcing strategies | | | |
| ln Sales | -0.024 | 0.033 | 0.159*** | |
| | (0.023) | (0.029) | (0.039) | |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | -0.053 | 0.073 | 0.356 | |
| - On $s^{(k)}$ | 0.010 | -0.007 | -0.010 | |
| Sample average of $s^{(k)}$ | 0.221 | 0.089 | 0.032 | |
| Observations | 11,690 | 11,690 | 11,690 | |
| Identified FE | 6.749 | 6.749 | 6.749 | |
| of which singletons | 4.081 | 4.081 | 4.081 | |
| | Most popular5-variety sourcing strategies | | | |
| ln Sales | 0.049 | 0.089* | 0.160*** | 0.145** |
| | (0.037) | (0.047) | (0.053) | (0.064) |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.105 | 0.191 | 0.342 | 0.310 |
| - On $s^{(k)}$ | -0.012 | -0.013 | -0.011 | -0.004 |
| Sample average of $s^{(k)}$ | 0.217 | 0.098 | 0.044 | 0.017 |
| Observations | 7.147 | 7.147 | 7.147 | 7.147 |
| Identified FE | 5.129 | 5.129 | 5.129 | 5.129 |
| of which singletons | 3.738 | 3.738 | 3.738 | 3.738 |

Notes: Regressions are estimated at the importer × product level. Regression samples include importers with exactly two, three, four or five varieties depending on the specification, observed from 2001 to 2006. For each NC8 product, only the most popular sourcing strategy is selected. Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). All regressions include fixed effects for years, four-digit industries and eight-digit products (× countries). Observations are weighted to give each firm an equal weight, independently of the number of imported products. The marginal impacts of an increase of sales by one standard deviation are computed at the mean of each sample in terms of sales, $s^{(1)}$, $s^{(2)}$, $s^{(3)}$, $s^{(4)}$ and $s^{(5)}$.

Table 15: Firm Size and the Intensive Margin of Trade: Sourcing-Strategy-Specific Coefficients

| | $\ln\left(\frac{s^{(1)}}{s^{(2)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(3)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(4)}}\right)$ | $\ln\left(\frac{s^{(1)}}{s^{(5)}}\right)$ |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | Firms sourcing (at least) from the 2 most popular countries | | | |
| ln Sales | 0.089*** | | | |
| | (0.005) | | | |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.225 | | | |
| - On $s^{(k)}$ | -0.032 | | | |
| Sample average of $s^{(k)}$ | 0.208 | | | |
| Observations | 210.341 | | | |
| Identified FE | 10.765 | | | |
| of which singletons | 1.371 | | | |
| | Firms sourcing (at least) from the 3 most popular countries | | | |
| ln Sales | 0.041*** | 0.127*** | | |
| | (0.006) | (0.009) | | |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.098 | 0.299 | | |
| - On $s^{(k)}$ | -0.012 | -0.021 | | |
| Sample average of $s^{(k)}$ | 0.218 | 0.091 | | |
| Observations | 79.453 | 79.453 | | |
| Identified FE | 9,180 | 9,180 | | |
| of which singletons | 2.107 | 2.107 | | |
| | Firms sourcing (at least) from the 4 most popular countries | | | |
| ln Sales | 0.014* | 0.058*** | 0.142*** | |
| | (0.008) | (0.010) | (0.013) | |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.031 | 0.130 | 0.319 | |
| - On $s^{(k)}$ | -0.001 | -0.010 | -0.013 | |
| Sample average of $s^{(k)}$ | 0.217 | 0.102 | 0.050 | |
| Observations | 40.614 | 40.614 | 40.614 | |
| Identified FE | 7.446 | 7.446 | 7.446 | |
| of which singletons | 2.299 | 2.299 | 2.299 | |
| | Firms sourcing (at least) from the 5 most popular countries | | | |
| ln Sales | 0.009 | 0.026* | 0.078*** | 0.151*** |
| | (0.010) | (0.013) | (0.016) | (0.020) |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | |
| - On dependent variable | 0.019 | 0.056 | 0.169 | 0.326 |
| - On $s^{(k)}$ | 0.000 | -0.004 | -0.008 | -0.008 |
| Sample average of $s^{(k)}$ | 0.212 | 0.107 | 0.058 | 0.031 |
| Observations | 23.989 | 23.989 | 23.989 | 23.989 |
| Identified FE | 5.967 | 5.967 | 5.967 | 5.967 |
| of which singletons | 2.219 | 2.219 | 2.219 | 2.219 |

Notes: Regressions are estimated at the importer × product level. Regression samples include importers with more than two, three, four or five varieties depending on the specification, observed from 2001 to 2006. For each NC8 product, only the most popular sourcing sub-strategy (of two, three, four or five countries respectively) is selected. Then, we restrict our sample to importers sourcing the considered product at least from this set of countries, while allowing disagreement about ranking of countries. Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). All regressions include fixed effects for years, four-digit industries and eight-digit products (× countries). Observations are weighted to give each firm an equal weight, independently of the number of imported products. The marginal impacts of an increase of sales by one standard deviation are computed at the mean of each sample in terms of sales, $s^{(1)}$, $s^{(2)}$, $s^{(3)}$, $s^{(4)}$ and $s^{(5)}$.

Table 16: Firm Size and the Intensive Margin of Trade, Exploiting the CES Structure

To make (47) operational, we adopt the following procedure. We first fix a number of varieties, $V$. For each product $k$, we select the $V$ varieties which appear in the highest number of sourcing strategies. We then keep all the firms that source from *at least* these $V$ countries - this is the key difference with the approach in Table 15. We then rank these $V$ countries for each firm and estimate the regression contained in (12) on the enlarged subsamples.[43] Table 16, which has the same structure as Table 15, contains the results. First of all note the usefulness of this approach in terms of increasing the sample size - for all cases, the number of observations increases by a factor of almost 3. This strengthens our earlier results considerably as (almost) all coefficients are positive and highly significant. As in Table 15, we also recover the monotonicity of the coefficients and the point estimates are very similar in magnitude. Hence, in Table 16 again, larger firms appear to concentrate their spending on their most preferred sourcing countries relative to smaller importers.

### B.2.6 The Domestic Expenditure Share

| Dep. Var.: | $\ln\left(\frac{s_D}{1-s_D}\right)$ | | | | |
|---|---|---|---|---|---|
| | Controlling for the sourcing strategy ($\Sigma$) | | | | Not contr. for $\Sigma$ |
| | (1) | (2) | (3) | (4) | (5) |
| ln Sales | 0.574*** | 0.658*** | 0.816*** | 0.733*** | -0.022*** |
| | (0.005) | (0.006) | (0.016) | (0.023) | (0.004) |
| ln Nb imported | -0.750*** | -0.831*** | - | - | - |
| products (8dig.) | (0.006) | (0.008) | | | |
| ln Nb countries | -0.679*** | - | - | - | - |
| | (0.009) | | | | |
| Fixed effects: | Years Industry | Years Industry Countries | Years Industry Products x Countries | Years Industry x Products x Countries | Years Industry |
| Impact of a 1 standard deviation increase in sales (from mean): | | | | | |
| - On dependent variable | 1.636 | 1.874 | 2.324 | 2.088 | -0.062 |
| - On $s_D$ | 0.224 | 0.239 | 0.261 | 0.250 | -0.013 |
| Sample average of $s_D$ | 0.699 | 0.699 | 0.699 | 0.699 | 0.699 |
| Observations | 160,103 | 160,103 | 160,103 | 160,103 | 160,103 |
| Identified FE | 334 | 59,181 | 143,873 | 148,947 | 334 |
| of which singletons | 3 | 51,222 | 137,721 | 143,650 | 3 |

Notes: Regressions at the importer level, observed from 2001 to 2006. Robust standard errors in parentheses with ***,** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include the following additional controls: a dummy indicating that the firm is an exporter, a dummy indicating that the firm is an affiliate of a larger corporate group, a dummy indicating that the latter has foreign affiliates (and not simply French domestic affiliates), a dummy indicating that the group has foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). Industry fixed effects are at the four digit level (the highest level available in the French classification of industries, which is slightly more detailed than the NACE). Product fixed effects are at the eight digit level (highest level of the EU Combined Nomenclature).

Table 17: Non-homothetic Import Demand: Firm Size and Domestic Shares

[43] This step is designed to rank the different varieties according to their price-adjusted quality. Recall that according to the theory the ranking of expenditure shares is indicative of the ranking of price adjusted qualities, for a given sourcing strategy. We use the firm-specific ranking to allow for disagreement in the ranking of these varieties across firms, as in the main text.