Firm Size and the Intensive Margin of Import Demand

Joaquin Blaum[1] & Claire Lelarge[2] & Michael Peters[3]

December 2017, WP #657

ABSTRACT

We use French microdata to test an ubiquitous property of firm-based models of importing. When firm efficiency is factor neutral and input prices and qualities are common across firms, firm size should have no effect on expenditure shares on the different products and varieties sourced, holding the extensive margin constant. We show that this property is not supported by the data. Holding the sourcing strategy fixed, we find that larger firms (i) have lower import shares, (ii) concentrate their import spending on their top varieties and (iii) pay higher prices for their imported inputs. Our findings imply that input trade, through the intensive margin, is less beneficial for larger firms. Our results are consistent with a complementarity between firm productivity and input quality.[4]

Keywords: Trade in Intermediate Inputs; Firm heterogeneity, Firm size, Non-homotheticities

JEL classification: F11, F12, F14, D21, D22, D24

NON-TECHNICAL SUMMARY

We analize the properties of a framework of importing which nests most of the contributions in the literature on input trade (e.g. Goldberg et al., 2010, Halpern et al., 2015, Gopinath and Neiman, 2014 or Antràs et al., 2017). Firms produce using a set of inputs that can be sourced domestically or from multiple foreign countries. The different varieties are imperfect substitutes in production and technology features constant returns to scale given the set of inputs sourced. Firms are price takers in input markets but can compete in output markets in an unrestricted way. Moreover, firm efficiency is factor neutral and input qualities and prices are common across firms. A prediction of this model is that, holding the set of sourcing countries fixed, firms' expenditure shares on the different input varieties depend only on the prices and qualities of the sourcing countries and should therefore be equalized across importing firms. Firm characteristics should be irrelevant: in this sense, the intensive margin of input trade is homothetic.

The empirical results however show systematic violations of this prediction. First, there are important idiosyncrasies in how firms allocate their spending across a common set of trading partners. To fix ideas, consider the case of parts of wrapping machinery, a particular 8-digit product. French importers who source this product from only Italy and Germany feature very different spending patterns. While 25% of these firms spend more than 90% of their import budget on the German variety, an equal number of firms has the exact opposite spending pattern and almost exclusively sources from Italian suppliers. We show that this pattern of disagreement is not specific to this particular product but holds generally for all products and sourcing strategies. Hence, firms' intensive margin sourcing patterns are by no means fully determined from country characteristics, as implied by standard models.

The Size Bias of Firms' Demand for Imported Varieties of Inputs
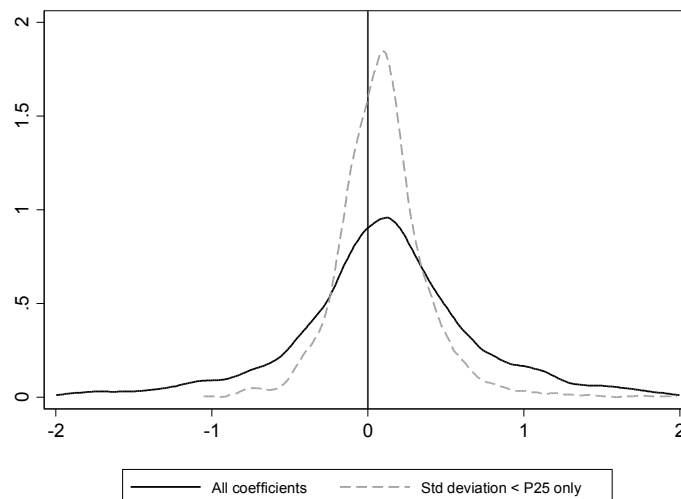
Notes: This figure represents the distribution of sourcing specific coefficients obtained from a regression of firms' expenditure shares on their main varieties (within imported products) on firms' sizes, as measured by their (ln) sales. Positive coefficients are indicative of larger firms biasing their expenditures towards their main varieties.

More importantly, we show that this disagreement across firms is not merely the result of idiosyncratic noise but is systematically related to firm size. First, large firms feature smaller import shares once the extensive margin is controlled for. This finding has normative implications as import shares summarize the productivity effects of

international sourcing (Blaum et al., 2016). In particular, this implies that, holding the extensive margin fixed, international trade is substantially less beneficial for larger firms.

Second, we find similar systematic rejections of the baseline model for the allocation of import expenditure across imported varieties (see *figure*). In particular, the distribution of foreign expenditure is substantially more concentrated for larger firms: we find that firm size is positively correlated with the share of spending allocated to the top country within narrowly defined products. More generally, larger firms feature a distribution of spending across foreign varieties which places higher mass on their higher ranked varieties.

Finally, we explore how the baseline framework could be amended to rationalize these facts. We provide two additional pieces of evidence that suggest the presence of a complementarity between firm productivity and input quality. First, we show that, holding the extensive margin fixed, larger firms pay higher prices for their inputs. This finding is consistent with Kugler and Verhoogen (2012), who interpret a similar correlation as reflecting a quality-productivity complementarity. Second, we show that larger firms pay particularly high prices for their top varieties. To the extent that prices reflect quality, these findings suggest that larger firms procure high quality inputs within foreign markets and they particularly do so in their top markets.

Comment interpréter la relation entre la taille des entreprises et la marge intensive de leur demande en intrants intermédiaires importés ?

Résumé

Dans cet article, nous nous appuyons sur des données micro-économiques d'entreprises françaises pour tester une propriété très générale des modèles d'import « à entreprises hétérogènes » proposés dans la littérature. Ces modèles reposent tous sur les hypothèses suivantes : la productivité des entreprises est supposée neutre par rapport aux facteurs de production, et les prix des intrants ainsi que leur qualité sont supposés identiques pour toutes les entreprises. Dans un tel cadre, il est possible de montrer qu'il ne devrait y avoir aucun lien entre la taille d'une entreprise et la décomposition de ses dépenses en intrants intermédiaires entre ses différentes variétés, *une fois contrôlé de la marge extensive* (c'est-à-dire de la liste des variétés d'intrants importés par l'entreprise). Cette propriété est cependant rejetée par les données. En contrôlant pour le choix des variétés importées, nous montrons que les entreprises de plus grande taille (i) présentent des parts de dépenses en intrants importés (dans le total des dépenses en consommations intermédiaires) plus faibles (ii) concentrent davantage leurs dépenses sur leurs principales variétés importées et (iii) paient des prix plus élevés pour leurs consommations intermédiaires importées. Ces résultats « à la marge intensive » impliquent que les entreprises de grande taille bénéficient relativement moins du commerce en biens intermédiaires que les entreprises de plus petite taille. Ils suggèrent de plus l'existence de complémentarités entre la productivité des entreprises et la qualité de leurs consommations intermédiaires.

Mots-clés : Commerce de biens intermédiaires, Hétérogénéité entre entreprises, taille des entreprises, demande non-homothétique.

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, providing new empirical regularities and building theoretical frameworks to account for them. About two-thirds of world trade, however, is accounted by trade in intermediate inputs. Additionally, imported inputs have been linked to firm productivity as well as to aggregate productivity during large crises. Understanding firms' behavior as importers is therefore important to account for global trade flows, as well as the normative implications of macroeconomic shocks and trade policies.

The literature has mostly relied on a common theoretical framework to study the import behavior of firms. Import demand stems from love-of-variety and quality effects, and is limited by the presence of fixed costs. While different contributions make different choices regarding the details of the environment, there is one theoretical implication that is common to essentially all frameworks in the literature: holding the set of trading partners fixed, firms' expenditure shares across trading partners should only depend on characteristics of the sourcing countries. In particular, firm characteristics such as size should not affect the allocation spending. This homotheticity property follows directly from the ubiquitous assumptions of factor neutral efficiency and common input prices and qualities across firms.

In this paper, we use French micro data to test this homotheticity prediction and show that it is not borne out in the data. Our empirical strategy exploits the distinction between the extensive and the intensive margin of importing. Standard theories of importing predict a positive relationship between firm size and import intensity. This relationship stems purely from firms' extensive margin of trade: in the presence of fixed costs, only sufficiently large firms import and larger firms import a larger set of varieties. These extensive margin patterns have been previously documented. In contrast, we focus on the intensive margin of import trade and present strong evidence against import demand being homothetic. We find that firms differ substantially in how they allocate their spending across a common set of trading partners. More importantly, these differences are systematically related to firm size. We find that firm size and import shares are negatively correlated, once the extensive margin is controlled for. That is, larger firms bias their spending towards domestic inputs for a given set of foreign trading partners. Larger firms also feature a more concentrated distribution of foreign expenditure. That is, they spend a larger proportion of their import budget on their most important foreign varieties, holding the sourcing strategy fixed.

We study a framework of importing which nests most of the contributions in the literature on input trade (e.g. Goldberg et al. (2010), Halpern et al. (2015), Gopinath and Neiman (2014) or Antràs et al. (2017)). Firms produce using a set of inputs that can be sourced domestically or from multiple foreign countries. The different varieties are imperfect substitutes in production and technology features constant returns to scale given the set of inputs sourced. Firms are price takers in input markets but can compete in output markets in an unrestricted way. Moreover, firm efficiency is factor neutral and input qualities and prices are common across firms. A prediction of this model is that, holding the set of sourcing countries fixed, firms' expenditure shares on the different input varieties depend only on the prices and qualities of the sourcing countries and should therefore be equalized across importing firms. Firm characteristics, notably firm efficiency, should be irrelevant. In this sense, the intensive margin of input trade is homothetic.

A crucial feature of our test is that we control for firms' extensive margin of trade. The theory stresses that the exact identity of the sourcing countries should be controlled for. In contrast, a common approach in the literature has been to rely on the number of sourcing countries as a sufficient statistic for the extensive

margin of trade, which is accurate only under special parameterizations of the model.[1] We show that this distinction is empirically important. In particular, we find substantial disagreement as to the identity of the sourcing countries among firms that source a given product from the same number of countries. That is, the number of countries turns out not to be a sufficient statistic for the sourcing set. We therefore perform our analysis by explicitly controlling for firms' sourcing strategies through sourcing-strategy-specific fixed effects or on subsets of firms that share exactly the same set of trading partners.

Our results show systematic violations of the benchmark importing framework. First, there are important idiosyncrasies in how firms allocate their spending across a common set of trading partners. To fix ideas, consider the case of parts of wrapping machinery, a particular 8-digit product. French importers who source this product from only Italy and Germany feature very different spending patterns. While 25% of these firms spend more than 90% of their import budget on the German variety, an equal number of firms has the exact opposite spending pattern and almost exclusively sources from Italian suppliers. We show that this pattern of disagreement is not specific to this particular product but holds generally for all products and sourcing strategies. Hence, firms' intensive margin sourcing patterns are by no means fully determined from country characteristics, as implied by standard models.

More importantly, we show that this disagreement across firms is not merely the result of idiosyncratic noise but is systematically related to firm size. With respect to domestic vs foreign spending, we find that large firms feature smaller import shares once the extensive margin is controlled for. In this way, the intensive margin operates in the opposite direction of the extensive margin. While large firms tend to have high import shares because they are active in more foreign markets, they actually have lower import shares once the extensive margin is kept constant. The former relationship is a form of non-homotheticity in import demand which stems from the extensive margin and has been previously documented - see e.g. Gopinath and Neiman (2014); the latter is a non-homotheticity that operates at the intensive margin and is the focus of this paper. Quantitatively, this intensive margin non-homotheticity is sufficiently strong to make domestic expenditure shares and firm size uncorrelated. In addition, these findings have normative implications as import shares summarize the productivity effects of international sourcing - see Blaum et al. (2016). Our results imply that, holding the extensive margin fixed, international trade is substantially less beneficial for larger firms.[2]

We find similar systematic rejections of the baseline model for the allocation of import expenditure across imported varieties. In particular, the distribution of foreign expenditure is substantially more concentrated for larger firms. We find that firm size is positively correlated with the share of spending allocated to the top country within narrowly defined products. More generally, larger firms feature a distribution of spending across foreign varieties which places higher mass on their higher ranked varieties.[3]

Finally, we explore how the baseline framework could be amended to rationalize these facts. We provide two additional pieces of evidence that suggest the presence of a complementarity between firm productivity and input quality. First, we show that, holding the extensive margin fixed, larger firms pay higher prices for their inputs. This finding is consistent with Kugler and Verhoogen (2012), who interpret a similar correlation as reflecting a quality-productivity complementarity. Second, we show that larger firms pay particularly high

---

[1]In a model with fixed costs to importing, the sourcing strategy can be summarized by the number of countries in the set when the fixed costs are constant across countries. This is the case in Halpern et al. (2015) and Gopinath and Neiman (2014). When fixed costs vary by country, as in Antràs et al. (2017), the identity of the sourcing countries may matter.

[2]This follows from the result in Blaum et al. (2016) that a high import share implies a large unit cost reduction associated with input trade. Our estimates imply that the gains from input trade are about three times larger for a firm with average sales relative to a firm with one standard deviation larger sales.

[3]Quantitatively, a one standard deviation increase in sales raises the expenditure share on the top foreign variety from 75% to 80%.

prices for their top varieties. To the extent that prices reflect quality, these findings suggest that larger firms procure high quality inputs within foreign markets and they particularly do so in their top markets. Furthermore, we show that the variation in input prices within countries is large relative to the cross-country variation. This finding is consistent with the large amount of disagreement among firms in the allocation of foreign expenditure across countries.

Related Literature. This paper is most closely related to a recent literature on firm-based models of importing. This literature takes a structural approach to study firms' importing decisions to explore a variety of questions. Examples include the relationship between firms' import activity and the associated productivity gains (see. e.g. Halpern et al. (2015), Goldberg et al. (2010), Ramanarayanan (2014)), the effect of changes in firms' import behavior on aggregate productivity (Gopinath and Neiman (2014)), the importance of country-specific fixed costs to shape firms' extensive margin responses to changes in trade policy (Antràs et al. (2017)), the dynamic process by which firms accumulate their foreign input suppliers (Lu et al. (2016)) or the importance of firms' import decisions to explain the exchange-rate disconnect (Amiti et al. (2014)).[4] Despite their different focus, all of these papers employ frameworks of importing where the import demand system is homothetic at the intensive margin. We therefore test a property of firms' import demand which is shared by many contributions in the literature.

One focus of our study is the relationship between firm size and the intensive margin of input trade. Several studies have empirically documented the relationship between firm size and the extensive margin of importing. In an early paper, Bernard et al. (2007) use US data to establish that importing is a rare activity and that importers are bigger and more productive than non-importers. More recently, several studies have shown that larger firms also import more products and from more countries- see Halpern et al. (2015) for Hungary, Gopinath and Neiman (2014) for Argentina, Antràs et al. (2017) for the US, or Eslava et al. (2015) for Colombia, among others.[5] In contrast, we document the effect of firm size on the intensive margin of importing. In particular, we show that larger firms feature smaller import shares once the extensive margin is controlled for. Because, as we have shown in Blaum et al. (2016), firms' domestic expenditure shares fully summarize the productivity consequences of importing, this implies that large firms benefit less from import trade than standard models would imply.

Finally, our paper is related to the quality-productivity complementarity hypothesis as stressed in Kugler and Verhoogen (2012). Using data from Colombia, they show that larger firms pay higher prices for their inputs and interpret this finding as reflecting a complementarity between firm productivity and input quality. Relatedly, Manova and Zhang (2012) show that exporters buy more expensive inputs from abroad. We document a similar association for the case of internationally traded inputs by French manufacturing firms: for a given supplying country, large firms pay higher prices for narrowly defined products.[6] Furthermore, we also show that large firms pay disproportionately high prices for their most important trading partners. This is also consistent with a complementarity between productivity and quality and with our finding that large firms bias their relative allocation of spending towards their top suppliers.

---

[4]Kasahara and Rodrigue (2008) and Amiti and Konings (2007) are early examples that empirically document the relationship between imported inputs and firm-level productivity.

[5]See also the recent surveys by Bernard et al. (2012) and Bernard et al. (2016) for a detailed overview of many empirical regularities regarding firms' participation in international markets.

[6]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 import status.

The structure of the paper is as follows. Section 2 contains a theory of import behavior and derives the intensive margin property which is the focus of the paper. Section 3 contains the empirical analysis using firms' expenditure shares. Section 4 provides additional evidence on unit values to suggest the presence of a complementarity between firm productivity and input quality. Section 5 concludes.

2   A Theory of Import Behavior

In this section, we present a general framework of importing that has been considered in the literature. Firms' import decisions are the solution to a static cost minimization problem. Foreign inputs imperfectly substitute domestic ones and may offer a quality or price advantage. Production is subject to constant returns, productivity is factor neutral and input prices and qualities are common across firms. Because we do not impose any assumptions on how the extensive margin of importing is determined, this framework nests many contributions in the literature, e.g. Halpern et al. (2015), Gopinath and Neiman (2014), Antràs et al. (2017), Blaum et al. (2016), Kasahara and Rodrigue (2008), Lu et al. (2016), Amiti et al. (2014) and Goldberg et al. (2010). We show that this setup implies that import demand is homothetic at the intensive margin. In particular, the distribution of spending across trading partners is independent of firm size.

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[7]:

$$
\begin{align}
y &= q\left(l, x; \varphi_i\right) \tag{1} \\
x &= f\left(x_1, .., x_n; \varphi_i\right) \tag{2} \\
x_k &= g_k\left(\{\eta_{cki} z_{ck}\}_{c \in \Sigma_{ki}}; \varphi_i\right), \tag{3}
\end{align}
$$

where l denotes primary factors, φ_i denotes firm efficiency, η_{cki} 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* Σ_i as the collection of sets of countries from which each product is sourced, i.e. $\Sigma_i \equiv \{\Sigma_{ki}\}_k$.

The main assumption encapsulated in the nesting structure of (1)-(3) concerns the degree of substitutability between different varieties. In particular, the marginal rate of substitution between two varieties of a product k is given by

$$
\frac{\partial f\left(x\right)/\partial z_{ck}}{\partial f\left(x\right)/\partial z_{c'k}} = \frac{\partial g_k\left(z\right)/\partial z_{ck}}{\partial g_k\left(z\right)/\partial z_{c'k}} \frac{\eta_{ck}}{\eta_{z_{c'k}}}, \tag{4}
$$

---

[7]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.

and hence does neither depend on the production functions q or f nor on any allocations $(x_{k'})$ or technologies $(g_{k'})$ in different product classes $k' \neq k$. While not crucial for our theoretical results, this feature will be convenient in the empirical analysis to test the theory at the product level.

We restrict the analysis to production functions f and $\{g_k\}_k$ with constant returns to scale *given the sourcing strategy* Σ.[8] This restriction 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 impose any restriction on how the sourcing strategy Σ 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.

Regarding the market structure, we assume that firms are price takers in input markets, i.e. they can purchase any quantities at given input prices $[p_{cki}]$. We assume that these prices contain all variable transport costs which accrue whenever a foreign variety is acquired. 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 two additional assumptions on technology and firm heterogeneity that are instrumental in deriving the key theoretical prediction of the paper. These assumptions are imposed in all contributions cited above. We consider technological differences across firms that are factor neutral and we restrict the degree of firm heterogeneity in the variable costs of inputs. Apart from efficiency φ_i, we allow firms to differ in their sourcing strategy Σ_i.[9]

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

$$
\begin{aligned}
f\left(x_1, .., x_n; \varphi_i\right) &= \varphi_i \hat{f}\left(x_1, .., x_n\right) \\
g_k\left(\{\eta_{cki} z_{ck}\}_{c \in \Sigma_{ki}}; \varphi_i\right) &= \hat{g}_k\left(\{\eta_{cki} z_{ck}\}_{c \in \Sigma_{ki}}\right)
\end{aligned}
$$

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

Assumption 2. *Input qualities and prices are common across firms. That is,*

$$
\eta_{cki} = \eta_{ck}, \, p_{cki} = p_{ck}, \, \forall c, k, i
$$

The framework laid out so far, which we refer to as the *standard model of importing*, is general enough to nest most of the existing theories of importing with firm heterogeneity in the literature. In particular, the literature typically considers the special case with a CES production structure of the form:

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

where m_{Fk} is a bundle of foreign varieties and D denotes the domestic input. (5) is a special case of the general technology in (1)-(3). In addition, virtually all contributions in the literature rely on fixed costs as a theory of the extensive margin. Our analysis does not rely on either of these specific assumptions in any way.

---

[8]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.

[9]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 fixed costs be firm-specific, as e.g. in Halpern et al. (2015) or Antràs et al. (2017).

2.2 The Intensive Margin of Importing

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

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

The function $\Gamma\left(\Sigma_i, x, \varphi_i\right)$ gives the cost of producing x units of the input composite for a firm with efficiency φ_i and sourcing strategy Σ_i. Denote by

$$\xi_{ck} \equiv \frac{\eta_{ck}}{p_{ck}} \tag{7}$$

the quality flow per dollar spent on product k imported from country c. For brevity, we refer to the $\xi's$ as qualities. The following proposition contains the central homotheticity property which is the focus of this paper.

Proposition 1. *Let $z_{ck}\left(\Sigma_i, x, \varphi_i\right)$ be the optimal amount sourced from country c of product k. The expenditure share of variety c in total spending for product k is given by*

$$s_{ck}\left(\Sigma_i, x, \varphi_i\right) \equiv \frac{p_{ck} z_{ck}(\Sigma_i, x, \varphi_i)}{\sum_{j \in \Sigma_{ki}} p_{jk} z_{jk}(\Sigma_i, x, \varphi_i)} = s_{ck}\left(\left[\xi_{ck}\right]_{c \in \Sigma_{ki}}\right), \tag{8}$$

that is, it depends only on the qualities of the varieties sourced within the product. Conditional on Σ_i, firm characteristics do not have an effect on the expenditure shares.

Proof. See Section A in the Appendix. □

Proposition 1 implies that, conditional on the sourcing strategy Σ, the within-product allocation of expenditure is equalized across firms. Firm characteristics, notably firm efficiency φ, have no effect on the expenditure shares other than through its effect on the set of varieties Σ_k. This result can be easily seen in the case of the canonical CES production function (see (5) above), where (8) reduces to the usual expression

$$s_{ck}\left(\Sigma_i, x, \varphi_i\right) = \frac{\xi_{ck}^{\rho-1}}{\sum_{j \in \Sigma_{ki}} \xi_{jk}^{\rho-1}}, \tag{9}$$

which directly shows that no firm characteristic determines the pattern of spending once Σ 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 φ. In this sense, Proposition 1 establishes that import demand is homothetic at the intensive margin. In Section 3 below, we test whether import demand actually satisfies this property using French micro data.

This result follows from Assumptions 1 and 2 above. That the intensive margin of import demand is homothetic, in the sense that the expenditure shares are independent of efficiency, follows from the assumptions that efficiency is factor neutral and production features constant returns. 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 (4). Together with the assumption that prices, qualities and variable

6

trade costs are common across firms, this property implies that expenditure shares do not depend on firm characteristics once the sourcing strategy is controlled for.[10]

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:

$$(\Sigma^*, y^*, l^*, x^*) = \arg \max_{(\Sigma, y, l, x)} \left\{ p(y) y - \Gamma(\Sigma_i, x, \varphi_i) - wl - \Omega(\Sigma) \right\}, \tag{10}$$

subject to (1), where $p(\cdot)$ denotes the demand that the firm faces, Γ 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[11] - see Gopinath and Neiman (2014), Halpern et al. (2015), Antràs et al. (2017) or Eslava et al. (2015). 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.[12] 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 Σ. 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).[13] We can ignore these difficulties because we focus on the homotheticity of import demand at the intensive margin, which holds regardless of the mechanics of the extensive margin.

3 Empirical Analysis

The goal of this section is to provide a test of Proposition 1, i.e. the prediction that the pattern of expenditure across varieties should not differ across firms once the extensive margin of trade is controlled for. Using data on French manufacturing firms, we show that this prediction is rejected, whatever the degree of flexibility of our empirical specification. We suggest a mechanism that can account for this result in Section 4.

---

[10]In fact, Halpern et al. (2015, p. 15) use this exclusion restriction explicitly as a form of identification. Precisely because expenditure shares on imported varieties should not depend on firm productivity holding the extensive margin of trade fixed, the cross-sectional variation in trade shares identifies firms' sourcing strategies.

[11]Formally, $\Omega(\Sigma) = \sum_{c \in \Sigma_k} \kappa_{ck}$ where κ_{ck} is the fixed cost required to import any quantity of product k from country c. Fixed costs are theoretically attractive as they allow to rationalize the fact that few firms import and most importers source from few countries.

[12]Note, however, that this does not imply that general results concerning the extensive margin cannot be derived. If for example the demand elasticity exceeds unity, more efficient firms import and more productive firms adopt a sourcing strategy that leads to lower unit costs, i.e. $\gamma(\Sigma(\varphi')) \leq \gamma(\Sigma(\varphi))$ if $\varphi' > \varphi$. Hence, similar to the exporting intuition, more productive firms sell more and thus have a higher incentive to reduce their marginal costs by incurring the fixed costs of importing additional products/varieties. However, again this does *not* imply that more efficient firms source *more* varieties or products. See Antràs et al. (2017) for a discussion.

[13]A common approach in the literature to ensure tractability of the extensive margin of importing problem is to assume that fixed costs are constant across countries. In this case, the imported inputs can be ranked along a single dimension, their price-adjusted quality - see Halpern et al. (2015), Gopinath and Neiman (2014) or Blaum et al. (2016). Antràs et al. (2017) provide a solution algorithm for the general case where fixed cost and qualities can vary in arbitrary ways.

3.1 Data

We start by providing a general overview of the dataset. A detailed description is contained in Appendix B. Because we are interested in the demand for inputs, we restrict the analysis to manufacturing firms. 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. Over the same period, French firms trade with a total of 230 to 240 countries depending on the year.[14]

		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 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. 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 raises to 86% for firms importing several varieties of a least one product.

3.2 Sourcing Strategies

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 Σ. Testing this prediction in the data requires taking a stand on the level of aggregation of products and varieties. We consider the maximum level of disaggregation and define a variety as an 8-digit product

---

[14]The French custom trade data are established in the "special trade system" (where overseas France is excluded from the statistical territory because of different custom restrictions).

coming from a particular country.[15] We now document two properties of firms' extensive margin of trade, defined at this level of disaggregation, which underscore the importance of appropriately controlling for the sourcing strategy.

Multi-Country Sourcing and Firm Heterogeneity. Two natural questions are the extent to which firms source a particular product, defined at the 8-digit level, from more than a single country and the extent of firm heterogeneity in these sourcing decisions. 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 that 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 one 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.[16] These statistics suggest that controlling for the extensive margin of trade is important, as there is meaningful heterogeneity in firms' sourcing strategies.
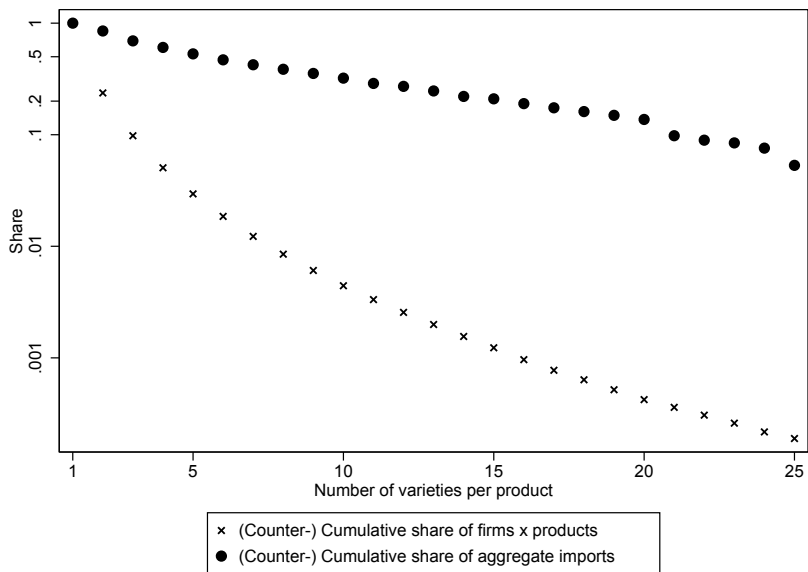
Notes: This figure shows the share of firms×products with at least C_{ik} varieties and the share of aggregate imports these firms×product pairs account for. We use 6 years of data from 2001-2006 and report the yearly average.

Figure 1: How Many Varieties Do Firms Source from Abroad?

Number vs Identity of Countries in Sourcing Set. Is the number of countries a sufficient statistic for the sourcing strategy? Several theoretical frameworks in the literature have imposed sufficient structure on

---

[15]In particular, given that we only focus on the time period between 2001 and 2006, we do not need to aggregate annual product categories into a time consistent classification as in Pierce and Schott (2012). This allows us to be as accurate as possible (with our data) in terms of the definition of products.

[16]In Section C in the Appendix, we report the distribution of spending across country varieties for multi-country firms. We find that multi-country firms spend a non-trivial amount of their total spending on their non-core varieties. On average, the most popular variety accounts for about $50\% - 60\%$ of a firm's import budget.

the import environment for this to be the case.[17] We now assess whether this property is born out by the data.

Figure 2 documents the extent to which firms that agree in the number of sourcing countries also agree in their identity. It shows that among firms sourcing a particular product from a given number of countries, there is substantial disagreement about the identity of these trading partners. To fix ideas, the left panel of Figure 2 focuses on a particular product as an example: "parts of packing or wrapping machinery, including heat-shrink wrapping machinery". We consider all firms that import this product from exactly two countries. We sort their sourcing strategies (two-country tupels) by popularity and depict the top ten sourcing strategies. If the number of countries was a sufficient statistic for firms' extensive margin of trade, this distribution would be degenerate. This is clearly not the case. While 30% of importers do source their parts of wrapping machinery from Germany and Italy, there are also about 20% of importers who pair the Italian variety with Belgium, Britain or Spain. Similarly, slightly more than 5% of firms have suppliers in China and supplement them with suppliers in Germany or Italy. Thus, while either Germany or Italy appear in all top ten sourcing strategies, there is ample disagreement as to the identity of the accompanying country.

(A) Product NC8 = 8422909 Sourcing Strategies with 2 Countries	(B) Averages Across All Products: Sourcing Strategies with 2 to 5 Countries
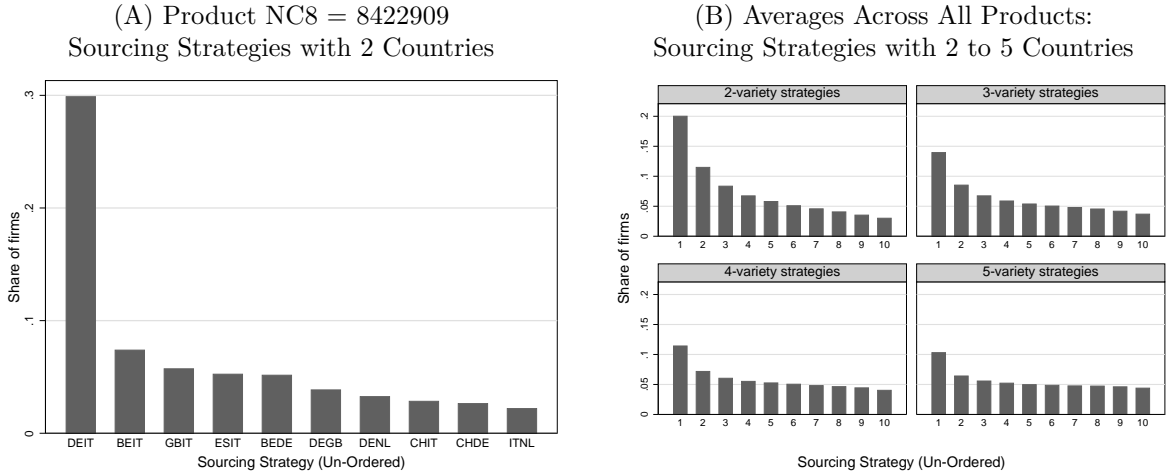
Notes: In the left panel we show the share of firms importing product NC8 = 84229090 ("parts of packing or wrapping machinery, including heat-shrink wrapping machinery") from exactly two countries by their respective sourcing strategies. In the right panel we show the share of firms by sourcing strategy 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 the yearly averages. Only products × cardinality of sourcing strategies with more than 10 observations are used in the computation.

Figure 2: Disagreement on the Extensive Margin

The right panel of Figure 2 shows that this pattern is not specific to this particular example but holds more generally. More specifically, we redo the analysis in 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.[18] We find considerable disagreement as to the countries from which firms

---

[17]See e.g. Gopinath and Neiman (2014), Halpern et al. (2015), Blaum et al. (2016) or Ramanarayanan (2014) who all assume, in the context of a model with fixed costs, that the per-country fixed cost of importing is constant across countries. This ensures that firms rank countries along a single dimension, their quality-adjusted price. Because firms agree in their ranking of countries by quality-adjusted price, as implied by Assumption 2, the number of countries perfectly summarizes the sourcing set. A notable exception is Antràs et al. (2017) who consider the case with country-specific fixed costs where the entire sourcing set is relevant.

[18]Consider for example the panel in the upper left corner. 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

import their products, even when 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 three countries, source this product from the most popular combination of countries. These patterns suggest that, to study the intensive margin of importing, we have to control for firms' entire sourcing set. While we also present results with the number of countries sourced, our preferred specification controls for firms' sourcing strategies in a non-restricted way.

3.3 The Intensive Margin I: Domestic vs. Foreign Expenditure

We start our analysis by focusing on a particular variety - the domestic one. Our focus on the French variety stems from the fact that firms' domestic expenditure shares contain important information to analyze the normative consequences of input trade. In particular, as we have shown in our earlier work in Blaum et al. (2016), the joint distribution of firm size and domestic expenditure shares fully summarizes the effect of input trade on consumer prices in a class firm-based models, which is nested in the framework of Section 2 above.[19]

Proposition 1 implies that the within-product domestic shares should be equalized across firms, once the set of trading partners within product, Σ_{ki}, is controlled for. Formally, $s_{Dki} = s_{Dk}\left(\Sigma_{ik}\right)$.[20] Empirically, we cannot look at this object directly. While we observe spending on foreign varieties at the 8-digit product level, our data does not contain information on domestic spending by product at this level of disaggregation. That is, we observe spending on French inputs at the firm but not at the firm $\times$ product level. However, Proposition 1 can be easily extended from the firm $\times$ product to the firm level. In particular, letting ν_{ki} denote firm i's expenditure share on product k, the theoretical structure above implies that $\nu_{ki} = \nu\left(\Sigma_i\right)$, i.e. ν_{ki} is fully determined by the entire sourcing strategy across all products. The share of domestic spending at the firm-level, s_{Di}, therefore satisfies

$$s_{Di} = \sum_{k=1}^{K} \nu_{ki}\left(\Sigma_i\right) \times s_{Dki}\left(\Sigma_{ki}\right) = s_D\left(\Sigma_i\right). \tag{12}$$

The sourcing strategy across all products, Σ_i, turns out to be a sufficient statistic for the domestic expenditure share. Therefore, the domestic shares should be equalized among firms that share a common Σ_i.

To study this implication empirically, we start by documenting how much of the variation in firms' domestic spending is accounted by the extensive margin of trade. Formally, we focus on the regression[21]

$$\ln\left(\frac{s_{Dit}}{1 - s_{Dit}}\right) = \alpha_S + \alpha_t + \alpha_\Sigma + u_{it}, \tag{13}$$

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.

[19]In particular, Blaum et al. (2016) show how the domestic expenditure shares can be used to measure the impact of a shock to the trading environment (e.g. a trade liberalization) on firms' production costs and the consumer price index. More specifically, the distribution of value added and changes in firms' domestic expenditure shares resulting from the shock can be used to recover the changes in firms' unit costs and the consumer price index. Furthermore, in the case of a reversal to input autarky, the data on domestic expenditure shares and value added are sufficient to measure the change in firms' unit costs and the consumer price index, regardless of whether Assumptions 1 and 2 above hold. However, studying counterfactuals other than autarky requires predicting the domestic expenditure shares after the shock, which entails resorting to a firm-based model of importing. The counterfactual domestic shares stemming from such model may depend on whether Assumptions 1 and 2 hold.

[20]The within-product domestic expenditure share of firm i, s_{Dki}, is defined as

$$s_{Dki} = \frac{p_{Dk}z_{Dk}}{\sum_{j \in \Sigma_{ki}} p_{jk}z_{jk}}, \tag{11}$$

where D denotes the domestic variety.

[21]We focus on the logit transformation of firms' domestic shares, i.e. $\ln\left(\frac{s_{Dit}}{1 - s_{Dit}}\right)$. This transformation turns out to be useful for the case of foreign varieties in Section 3.4. For symmetry we also focus on this transformation here.

where α_S and α_t are industry and year fixed effects and α_Σ denotes the control for firm i's sourcing strategy. We are interested in the residual variation u_{it}. Table 2 contains the results for different choices regarding α_Σ. We report the results after dropping any sourcing strategy that only has a single observation. We focus on this sample without singletons because this is where we can test the theory.[22]

Column 1 reports the variation in domestic shares without any controls for the extensive margin. The variation across years and industries accounts for about 10% of the cross-firm variation in domestic expenditure shares. Columns 2 and 3 add a full set of product *set* or country *set* fixed effects, respectively. That is, we look at the set of firms sourcing the exact same 8-digit products from abroad (column 2) or import from the exact same set of countries (column 3). In the case of products (countries), this amounts to controlling for about 7,000 (8,000) fixed effects in the corresponding samples of 41,745 (116,125) observations.[23] Controlling for the sourcing strategy in this way increases the R^2 of the regression to 44% and 28%. Conversely, between 56% and 72% of the variation remains unexplained. Column 4 includes both set of fixed effects. The results are similar to the "products only" specification in column 2. Finally, column 5 controls for the entire sourcing strategy, i.e. the whole set of interactions of product set and county set fixed effects, as prescribed by the theory. Even then, we find that one third of the cross-sectional variation in domestic shares remains unexplained. Finally, for completeness, column 6 allows firms' domestic shares to vary across industries within each sourcing strategy. This still leaves one quarter of the variation unexplained.

| | | | $\ln\left(\frac{s_D}{1-s_D}\right)$ | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Observations	168,662	168,662	168,662	168,662	168,662	168,662
Identified FE	334	134,768	60,743	137,607	150,317	156,152
of which singletons	3	127,917	52,537	130,466	143,572	150,307
Variance decomposition, removing singletons (Correia, 2015; Gaure, 2013):						
R^2	0.105	0.444	0.280	0.487	0.642	0.762
Year and ind. FE	0.105	0.124	0.087	0.126	0.163	0.000
Product and Sourcing Strat. FE	0.000	0.401	0.207	0.447	0.612	0.772
Covariance term	0.000	-0.081	-0.014	-0.086	-0.133	-0.010
Fixed effects:	Years	Years	Years	Years	Years	Years
	Industry	Industry	Industry	Industry	Industry	Industry $\times$
		Products		Products	Products $\times$	Products $\times$
			Countries	Countries	Countries	Countries

Notes: Regressions at the importer level, observed from 2001 to 2006. 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). The row entitled "Year and Industry Fixed Effects" contains the contribution to total variance of year and industry dummies, except in the fully saturated specification of column (6), where industry dummies are interacted with product and sourcing strategies. The row entitled "Product and Sourcing Strategy Fixed Effects" contains the contribution to total variance of product and / or sourcing strategy dummies or product $\times$ sourcing strategy interaction dummies.

Table 2: Variance Decomposition for Domestic Shares

Figure 3 displays these results graphically. In particular, it plots the distribution of residuals from the regressions in Table 2. For brevity, we report the specifications of columns 1 and 5 only, which we interpret as the unconditional variation in domestic shares and the variation conditional on the extensive margin of trade, respectively. The theory of Section 2 implies that the conditional distribution should be degenerate. Figure 3 shows that this is clearly not the case. While the mass in the tails of the distribution shrinks considerably,

---

[22]Singletons are observations whose sourcing strategy is unique. It follows that Proposition 1 cannot be tested on these observations. See Correia (2015) and Gaure (2013) for a discussion of singletons in regression settings with fixed effects.

[23]These numbers are computed as 134,768 - 127,917 and 60,743 - 52,537 respectively. Samples without singletons contain 168,662 - 127,917 = 41,745 and 168,662 - 52,537 = 116,125 observations, respectively.

there is still ample heterogeneity in the extent to which firms rely on domestic inputs within narrowly defined sourcing strategy bins. This fact is inconsistent with the vast majority of firm-based models of the literature.
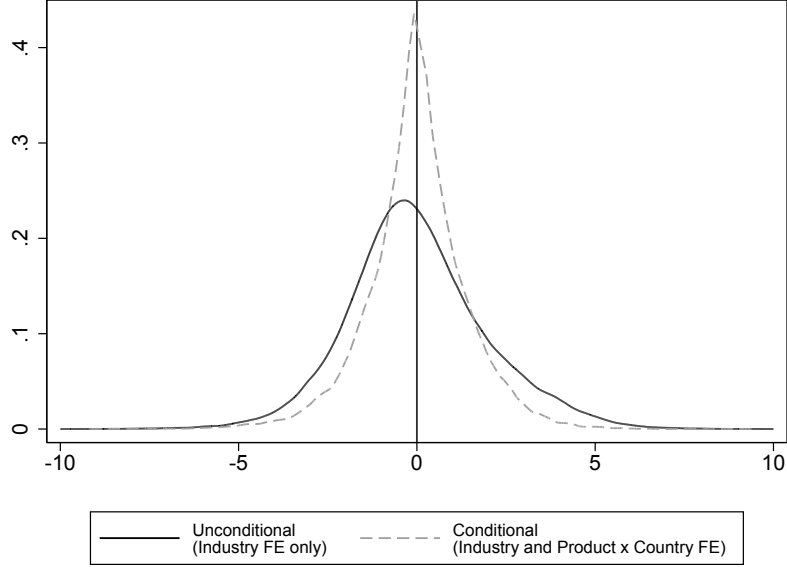
Notes: This figure plots the residuals obtained for specifications (1) and (5) in Table 2. "Singletons" (observations that are unique in at least one dimension of the fixed effects, and whose associated residual is therefore nil) are removed (Correia, 2015; Gaure, 2013).

Figure 3: Residuals in the Variance Decomposition of Domestic Shares

While the results of Table 2 and Figure 3 show that firms' extensive margin of trade is not a sufficient statistic for the allocation of expenditure, we now explore whether this variation is systematic and can be linked to particular firm characteristics. We are particularly interested in the effect of firm size, as we know from Blaum et al. (2016) that the joint distribution of firm size and domestic shares determines the welfare consequences of input trade. Equation (12) implies the exclusion restriction $s_{Di}(\Sigma_i, \varphi_i) = s_{Di}(\Sigma_i)$. That is, conditional on firms' sourcing set, firm productivity should be uncorrelated with domestic expenditure shares. Through the lens of the theory of Section 2, this implies that firm size should be uncorrelated with domestic shares. To test this restriction, we augment the specification in (13) by including firm sales as a measure of firm size:

$$\ln\left(\frac{s_{Dit}}{1 - s_{Dit}}\right) = \alpha_S + \alpha_t + \alpha_\Sigma + \beta \times \ln(S_{it}) + X'_{it}\mu + u_{it}, \tag{14}$$

where the fixed effects α are defined as above and $\ln(S_{it})$ denotes log sales. In addition, we also include additional firm-level characteristics in X_{it} to control for differences in technology and the input market environment within industries.[24] The main parameter of interest is β, which - according to the theory - should be zero.

The results are reported in Table 3. The different columns control for firms' sourcing strategy in different ways. While column 2 controls linearly for the cardinality of the sourcing set, i.e. the number of products sourced and the number of countries imported from, column 5 contains the most stringent specification

---

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

Dep. Var.:			$\ln\left(\frac{s_D}{1-s_D}\right)$		
	(1)	(2)	(3)	(4)	(5)
ln Sales	-0.022***	0.574***	0.658***	0.816***	0.733***
	(0.004)	(0.005)	(0.006)	(0.016)	(0.023)
ln Nb imported	-	-0.750***	-0.831***	-	-
products (8 dig.)		(0.006)	(0.008)		
ln Nb countries	-	-0.679***	-	-	-
		(0.009)			
Fixed effects:	Years	Years	Years	Years	Years
	Industry	Industry	Industry	Industry	Industry $\times$
			Countries	Products $\times$	Products $\times$
				Countries	Countries
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	-0.062	1.636	1.874	2.324	2.088
- On s_D	-0.013	0.224	0.239	0.261	0.250
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	334	59,181	143,873	148,947
of which singletons	3	3	51,222	137,721	143,650

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 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 3: Firm Size and Domestic Shares

that controls for all interactions of industry, product set and country set fixed effects. Regardless of the specification, Table 3 shows that we can comfortably reject the restriction in (12): holding the sourcing strategy fixed, firm size is a significant predictor of the reliance on domestic inputs.

The particular direction in which the theory is rejected is also economically meaningful. The coefficients on sales in Table 3 are all positive, i.e. larger firms spend a *larger* fraction of their resources domestically. This relationship is in sharp contrast to the effect that firm size has on the domestic shares 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.[25] 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, which has strong empirical support, implies a *negative* relation between firm size and the domestic share, through the number of varieties imported. This is what we find, as both the number of foreign countries and products negatively correlate with the domestic share - see column 2 in Table 3. More importantly, however, Table 3 also shows that, 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. (2016) that this unconditional correlation is essentially zero for the sample of importers.

The results in Table 3 imply that, holding the extensive margin of trade constant, international trade

---

[25]For the positive relation between firm size and the number of imported varieties see Gopinath and Neiman (2014) for Argentina, Halpern et al. (2015) for Hungary, or Blaum et al. (2016) for France.

is *less* beneficial for larger firms. This follows from the result in Blaum et al. (2016) that a low domestic share implies a large unit cost reduction associated with input trade. Quantitatively, the estimates of Table 3 imply that a one standard deviation increase in sales (from sample means) increases the expenditure share on domestic varieties by around 20 percentage points, from about 70% to 90%. Through the lens of the sufficiency result in Blaum et al. (2016), these estimates imply that the gains from input trade are 17% for a firm with average sales vs. 5% for a firm with one standard deviation larger sales.[26]

3.4   The Intensive Margin II: The Allocation of Spending Across Foreign Varieties

Foreign varieties provide another dimension in which to test the theory. An empirical advantage is that we observe foreign expenditure *by product.* In contrast to the analysis of the domestic vs. foreign expenditure, which was performed at the firm-level, we now study the allocation of spending across foreign varieties at the firm-product level.[27]

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 Σ_k should have the exact same expenditure shares on each supplying country $c \in \Sigma_k$. We now start to investigate the extent to which this result is borne by the data. 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.[28] In the left panel of Figure 4, we depict the distribution of the expenditure share on the German variety. Under assumptions 1 and 2, 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. Additionally, the distribution has substantial mass on the extremes. Thus, even among firms who import a particular product from Germany and Italy only, there tends to be a dominant supplier, which accounts for more than 90% of the firm's import budget. The identity of this supplier, however, varies across firms. While 25% of firms almost exclusively source from Germany, an almost equal number of firms rely mostly on the Italian variety.

In the right panel of Figure 4, we consider all products and sourcing strategies and show that the pattern found in the previous example holds more generally. In particular, for each product and sourcing strategy, we define the most popular variety as the one which most firms have as their dominant supplier (i.e. with the largest expenditure share) and compute the dispersion (as measured by the interquartile range) of expenditure shares on this variety. We then average these dispersions across products and sourcing strategies of a given cardinality. For example, for all products × sourcing strategies combinations involving exactly two countries,

---

[26]Blaum et al. (2016) show that the gains from input trade, defined as the change in unit cost between autarky and the current equilibrium, holding general equilibrium prices fixed, are given by

$$\ln\left(u_i^{aut}/u_i\right) = \frac{\gamma}{1 - \varepsilon} \times \ln(s_{Di}),$$

where u_i, u_i^{aut} are the unit cost in the current equilibrium and autarky, respectively, γ is the share of materials in value added, and ε is the elasticity of substitution between domestic and foreign varieties. On a similar sample of French firms Blaum et al. (2016) estimate $\gamma = 0.61$ and $\varepsilon = 2.38$.

[27]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.

[28]The specific product NC8 84229090 has been selected for its popularity, which allows us to plot Figure 4 without violating statistical secrecy.

(A) Product NC8 = 8422909
Sourcing Strategy = {DE, IT}

(B) Averages Across
All Products × Sourcing Strategies
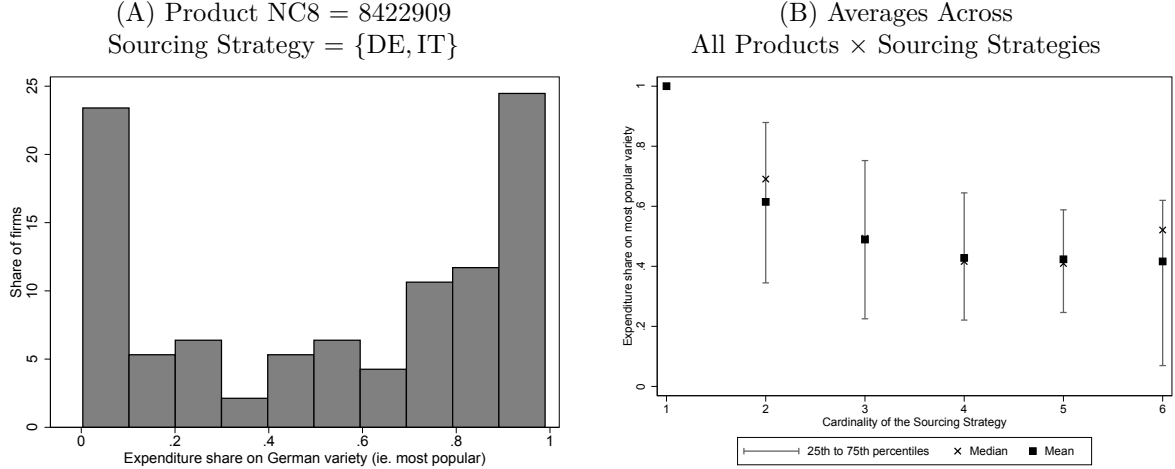
Notes: The left panel shows the dispersion in the share of expenditures of "parts of packing or wrapping machinery (including heat-shrink wrapping machinery)" from Germany, for firms sourcing this product from Germany and Italy, as of 2002. The right panel depicts moments of the distribution of the expenditure share on the most popular variety within each product×sourcing strategy, depending on the cardinality of the (un-ordered) sourcing strategy. In both panels, we use 6 years of data from 2001-2006 and report the yearly averages. Only products×sourcing strategies with more than 5 observations are used in the computation.

Figure 4: Disagreement on the Intensive Margin: Main Varieties within Sourcing Strategies

the interquartile range of the expenditure share on the most popular variety ranges from 40% to 90%. Proposition 1 implies that there should be no dispersion and *all* firms should spend exactly 60% (the mean) in the most popular variety. Figure 4 therefore indicates substantial variation in how firms allocate their foreign expenditure holding the extensive margin of trade constant.[29]

As we did for the case of the domestic variety, we formalize this notion of disagreement through a variance decomposition exercise. Table 4 reports the R^2 of a regression of firms' expenditure shares on their most important trading partner *within* a given 8-digit product on a full set of sourcing strategy and industry fixed effects. Performing the analysis at the firm × product level increases our sample size substantially relative to the case of the domestic variety reported in Table 2 above. Thus, we can control for the sourcing strategy in the most stringent way, i.e. via the interaction of product fixed effects and the set of supplying countries, and still have ample degrees of freedom.

Table 4 confirms the findings for the domestic share in Table 2: there is 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).[30] 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 4 - 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 4 and Table 4 indicate that firms' sourcing strategies are not a sufficient statistic for the distribution of foreign spending as predicted by standard models of importing.

---

[29]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.

[30]To be precise, the sample without singletons in column 1 of Table 4 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.

Dependent variable:			$\ln\left(\frac{s_1}{1-s_1}\right)$		
Nb. varieties:	≥ 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 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 4: Within Sourcing-Strategy Variance of Firms' Expenditure Shares

The Role of Firm Size. The disagreement on the intensive margin documented above could emerge from purely idiosyncratic differences across firms, e.g. in perceived country quality. For the product considered in the left panel of Figure 4, this would amount to some firms perceiving a higher quality from the German variety, while others find the Italian one to be superior. In contrast, there could be more systematic forces at work if the differences in the pattern of foreign spending found above are linked to particular firm characteristics. As we did for the domestic variety, we now explore the role of firm size.

To abstract from idiosyncratic quality perceptions, we focus on the distribution of foreign spending across ranked varieties - i.e. the firm's most important supplier, the second most important, etc. That is, for a given set of supplying countries, we abstract from the identity of the supplying countries when calculating firms' expenditure shares. An advantage of this approach is that it allows for comparisons across different product × sourcing strategies, as e.g. the most important supplier is always defined. In this way, we can study the relation between firm size and the distribution of foreign spending systematically across all product-supplier pairs. Additionally, characterizing the distribution of firm i's spending through order statistics, i.e. the ranked expenditure shares, also has economic content as the ranking of firms' expenditure shares reflects the ranking of price adjusted qualities ξ. To see this, consider the case where the foreign varieties within products are combined in a CES way as in (5). In this case, the share of expenditure allocated to variety c of product k by firm i is given by (9), which directly shows how expenditure shares depend on price-adjusted qualities.

More generally, Proposition 1 implies that

$$s_{jk}^i = s_{jk}\left(\Sigma_{ki},\varphi_i\right) = s_{jk}\left(\Sigma_{ki}\right) \text{ for all } j = 1,...,|\Sigma_{ki}|, \tag{15}$$

where, with a slight abuse of notation, s_{jk}^i denotes the j-th order statistic of firms i's expenditure shares on product k and $|\Sigma_{ki}|$ is the total number of countries sourced by firm i of product k.[31] The exclusion restriction in (15) implies that firm productivity should not be correlated with spending patterns on the

---

[31] While (15) is a necessary condition for (8), it is not sufficient. The reason is that (8) concerns the distribution of spending across the set of countries including their identities c. Consider for example a set of firms sourcing product k from two countries A and B. While (8) implies that all firms should spend the same share of their import budget on country A, the exclusion restriction above only implies that all firms should have the same allocation of spending across the two countries, irrespective of their identity.

intensive margin. More productive firms should not bias their spending towards inputs of higher, nor lower, price-adjusted qualities.

We start our analysis by focusing on firms' most important variety, which on average accounts for about 50% to 60% of firms' import budget (see Section C in the Appendix). We focus on the specification

$$\ln\left(\frac{s_{1kt}^i}{1 - s_{1kt}^i}\right) = \alpha_S + \alpha_t + \alpha_k + \alpha_{\Sigma_{ki}} + \beta \times \ln(S_{it}) + X_{it}'\mu + u_{ikt}, \tag{16}$$

where s_{1kt}^i is firm i's expenditure share on its top variety and α_S, α_t and α_k are sector, time and product fixed effects. As above, $\alpha_{\Sigma_{ki}}$ denotes our control for firms' sourcing strategy, which is now productivity specific,[32] and $\ln(S_{it})$ and X_{it} denote log sales and the same firm-specific covariates as employed in Table 3.

In a second set of regressions, we focus on the entire distribution of expenditure shares, relative to the top variety, according to the following specification[33]:

$$\ln\left(\frac{s_{1kt}^i}{s_{jkt}^i}\right) = \alpha_S + \alpha_t + \alpha_k + \alpha_{\Sigma_{ki}} + \beta \times \ln(S_{it}) + X_{it}'\mu + u_{ikt}, \text{ for } j = 2, 3, ., , |\Sigma_{ki}|. \tag{17}$$

While (16) asks whether larger firms have systematically different expenditures shares on their most important trading partner, (17) asks whether the distribution of spending across different varieties is correlated systematically with firm size. According to Proposition 1, the distribution of spending should not depend on firm size, i.e. β in (16) and (17) should be zero in all specifications.

The results are contained in Table 5. The first column focuses on the top variety, following (16), while the remaining columns focus on the entire distribution of expenditure shares as in (17). We report results for the cases of up to five varieties.[34] As with domestic shares, we distinguish two ways to control for the extensive margin of trade. In panel A, the sourcing strategy of product k, Σ_{ki}, is proxied by its cardinality, $\ln|\Sigma_{ki}|$, i.e. the number countries product k is sourced from. In Panel B, we control for the sourcing strategy by including a complete set of industry × product × sourcing-strategy fixed effects, $\alpha_{\Sigma_{ki}}$, as required by the theory. That is, we compare firms in the same industry who source their inputs from exactly the same set of countries within 8-digit products.

The results show that we can reject the homotheticity property implied by Proposition 1. We find that the estimated coefficient for sales is highly statistically significant and of a similar order of magnitude regardless

---

[32]In the CES case, (16) is given by

$$\ln\left(\frac{s_{1kt}^i}{1 - s_{1kt}^i}\right) = \underbrace{\ln\left(\xi_{1,\Sigma_{ki}}^{\rho-1}\right) - \ln\left(\sum_{j=2}^{|\Sigma_{ki}|} \xi_{j,\Sigma_{ki}}^{\rho-1}\right)}_{\alpha_{\Sigma_{ki}}} + \beta \times \ln(S_i) + X_i'\mu + u_{ik},$$

where ξ_{j,Σ_k} is the j-th order statistic of $[\xi_{ck}]_{c\in\Sigma_k}$.

[33]In the CES case, (17) is given by

$$\ln\left(\frac{s_{1kt}^i}{s_{jkt}^i}\right) = \underbrace{(\rho - 1) \times \ln\left(\frac{\xi_{1k}}{\xi_{jk}}\right)}_{\alpha_k} + \beta \times \ln(S_i) + X_i'\mu + u_{ik}.$$

Note that controlling for the sourcing strategy is not required in this case. We exploit this property in Section C.3 in the Appendix to increase the statistical power of our test.

[34]For sourcing strategies involving more than five countries (per 8 digit product), we lack statistical power to test (17) as there are too few observations within sourcing strategies. Note that the sample size decreases substantially as we consider less important varieties, since the number of firms sourcing at least three, four or five foreign varieties is small - see Figure 1.

Panel (A): Controlling for the Number of Countries

	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
ln Sales	0.114***	0.107***	0.148***	0.174***	0.185***
	(0.006)	(0.006)	(0.012)	(0.018)	(0.026)
ln Nb countries	-1.446***	-0.879***	-1.540***	-2.045***	-2.493***
	(0.015)	(0.015)	(0.030)	(0.051)	(0.079)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.304	0.286	0.410	0.455	0.474
- On s_k	0.052	-0.041	-0.023	-0.014	-0.008
Sample average of s_k	0.756	0.206	0.080	0.041	0.024
Observations	739,167	739,167	306,573	158,652	92,722
Identified FE	462,159	462,159	226,593	126,607	77,729
of which singletons	339,672	339,672	181,470	105,941	67,223

Panel (B): Controlling for the Entire Sourcing Strategy

	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
ln Sales	0.122***	0.119***	0.178***	0.202***	0.179***
	(0.008)	(0.008)	(0.019)	(0.039)	(0.064)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.324	0.319	0.451	0.491	0.418
- On s_k	0.055	-0.045	-0.026	-0.015	-0.008
Sample average of s_k	0.756	0.206	0.080	0.041	0.024
Observations	739,167	739,167	306,573	158,652	92,722
Identified FE	534,061	534,061	261,019	144,718	87,588
of which singletons	435,837	435,837	232,960	134,865	83,661

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. 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). Regressions in panel (A) include fixed effects for years, 4-digit industries interacted with 8-digit products and sourcing countries (of the specific varieties considered). Regressions in panel (B) include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product). 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 5: Firm Size and the Intensive Margin of Trade

of how the sourcing strategy is controlled for. Importantly, the coefficient is strongly positive: larger firms bias their expenditure towards their most important foreign varieties and against the minor ones, holding the set of countries constant. Note that these results are consistent with the findings on domestic spending of Table 3 as the domestic variety tends to account for most of the material expenditure. Note also that the coefficients tend to increase in magnitude with the order of the statistic from columns 2 to 5. This implies 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 five percentage points, i.e. from about 75% to 80%.[35,36,37]

In the above specifications, we pool all importers sourcing at least V varieties, for V ranging between two and five. For instance, column 1 in Table 5 contains firms sourcing two or more varieties. While we control for sourcing strategy fixed effects, the coefficient of interest β is identified by pooling firms with different sourcing strategies. To show that this does not drive our results, we report a more demanding empirical specification in Table 6. We consider the same specification but 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 specifications in (16) and (17) pooling all firm × 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.[38]

The results, shown in Table 6, are very similar to the ones of Table 5, 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. A limitation of this exercise is that the sample size is small relative to the number of products (and the associated fixed effects) in all panels. This reduces the statistical power of our test. In Section C.3 in the Appendix, we therefore impose the additional assumption that technologies are CES. This allows us to substantially increase our sample sizes. Using this approach, we also confirm the results obtained in Table 5 and find that all coefficients (but one) are significantly positive.

---

[35]These marginal effects are computed by inverting the logit function to fully take account of its high non-convexity. Formally, let $D_1 = \ln\left(\frac{s_1}{1-s_1}\right)$ and $D_{1k} = \ln\left(\frac{s_1}{s_k}\right)$ such that $ds_1 = \frac{e^{D_1+dD_1}}{1+e^{D_1+dD_1}} - \frac{e^{D_1}}{1+e^{D_1}}$ and $ds_k = \frac{s_1+ds_1}{e^{D_{1k}+dD_{1k}}} - s_k$. These quantities are evaluated at the sample means of s_1 and s_k in each column of Table 5. Note that (16) and (17) concern the distribution of *import* spending. As we showed above in Table 3, firm size is negatively correlated with the import share. The unconditional expenditure share on foreign country c for product k is therefore given by $s_{ck}^U = s_{ck} \times (1 - s_{D,k})$. Hence, the change in total spending on variety c on product k is given by $ds_{ck}^U = ds_{ck} \times (1 - s_{D,k}) - s_{ck} \times ds_{D,k}$. As pointed above, our data does not contain information on domestic spending at the product level. If the domestic share was equalized across products and hence equal to the firm-level average of 0.625 (see column 8 of Table 3), the results in panel (B) of Table 3 translate into average shares of expenditure allocated to the main, second, third, fourth or fifth imported of varieties of 0.283, 0.077, 0.030, 0.015 and 0.009 respectively. The corresponding marginal impacts of a one standard deviation increase in sales are all negative: -0.156, -0.052, -0.022, -0.011 and -0.006 respectively. Hence, even though the share of import spending on the top variety increases, the unconditional share declines as larger firms also increase their reliance on domestic inputs.

[36]Reassuringly, we also obtain that the cumulated negative impact of a one standard deviation increase in sales on minor varieties broadly matches the positive impact on the main variety. This was not guaranteed to happen as the coefficients in the different columns of Table 5 are estimated from different samples. Given the skewness of the distribution of expenditure shares earlier described in Figure 6, this implies that the relative impact of such a perturbation on minor varieties is relatively much larger than the impact on the main variety (around 20%-30% against ca. 7%).

[37]In Appendix C.2 we show that the obtained results are quantitatively very similar with employment or TFP as alternative proxies for firm efficiency once we control for technological heterogeneity across firms. We also show that the results are robust to controlling for firm fixed effects.

[38]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.

| | $\ln\left(\frac{s_1}{1-s_1}\right)$ | $\ln\left(\frac{s_1}{s_2}\right)$ | $\ln\left(\frac{s_1}{s_3}\right)$ | $\ln\left(\frac{s_1}{s_4}\right)$ | $\ln\left(\frac{s_1}{s_5}\right)$ |
| | (1) | (2) | (3) | (4) | (5) |

Most popular 2-variety sourcing strategies					
ln Sales	0.091***	0.091***			
	(0.007)	(0.007)			
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.225	0.225			
- On s_k	0.033	-0.033			
Sample average of sk	0.801	0.199			
Observations	83,778	83,778			
Identified FE	10,547	10,547			
of which singletons	2,189	2,189			

Most popular 3-variety sourcing strategies					
ln Sales	0.072***	0.058***	0.177***		
	(0.013)	(0.013)	(0.018)		
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.168	0.135	0.414		
- On s_k	0.033	-0.019	-0.022		
Sample average of s_k	0.712	0.218	0.070		
Observations	24,747	24,747	24,747		
Identified FE	8,740	8,740	8,740		
of which singletons	3,767	3,767	3,767		

Most popular 4-variety sourcing strategies					
ln Sales	-0.011	-0.024	0.033	0.159***	
	(0.022)	(0.023)	(0.029)	(0.039)	
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	-0.024	-0.053	0.073	0.356	
- On s_k	-0.005	0.010	-0.007	-0.010	
Sample average of s_k	0.658	0.221	0.089	0.032	
Observations	11,690	11,690	11,690	11,690	
Identified FE	6,749	6,749	6,749	6,749	
of which singletons	4,081	4,081	4,081	4,081	

Most popular 5-variety sourcing strategies					
ln Sales	0.065*	0.049	0.089*	0.160***	0.145**
	(0.036)	(0.037)	(0.047)	(0.053)	(0.064)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.139	0.105	0.191	0.342	0.310
- On s_k	0.032	-0.012	-0.013	-0.011	-0.004
Sample average of s_k	0.625	0.217	0.098	0.044	0.017
Observations	7,147	7,147	7,147	7,147	7,147
Identified FE	5,129	5,129	5,129	5,129	5,129
of which singletons	3,738	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 6: Firm Size and the Intensive Margin of Trade: Sourcing-Strategy-Specific Coefficients

Finally, we extend the analysis beyond the most popular sourcing strategy and consider all sourcing strategies. This amounts to allowing the β coefficients in Equation (16) to vary by sourcing strategy. In practice, this requires to estimate Equation (16) on subsamples of firm × product pairs that share exactly the same sourcing strategy and yields a distribution of estimated β's across product-specific sourcing strategies. More specifically, let σ_k be a particular sourcing strategy for product k. We then consider the following specification for each (k, σ_k)[39]:

$$\ln\left(\frac{s_{1kt}^i}{1 - s_{1kt}^i}\right) = \alpha_{tk} + \beta_{\sigma_k} \times \ln(S_{it}) + X'_{it}\mu_k + u_{ikt}, \text{ for all } k \text{ and } \sigma_k \tag{18}$$

In this equation, β_{σ_k} denotes the coefficients associated with log sales for the particular sourcing strategy σ_k for product k, α_{tk} are product × year fixed effects and X_{it} are again the usual control variables, whose effect on the expenditure share on the top variety we allow to vary by product. We estimate (18) for all products × sourcing strategies σ_k which have at least 10 observations in order to insure a minimal amount of statistical power. This leaves us with 6,751 coefficients β_{σ_k}, which we depict in Figure 5.
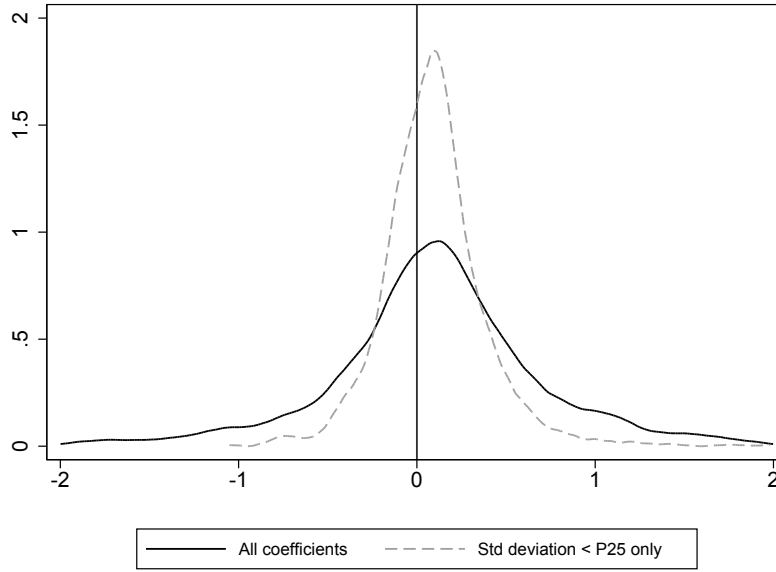
Notes: This figure represents the distribution of sourcing specific coefficients obtained from the regression (18). Only sourcing strategies × products with more than 10 observations are included into the sample. We consider the years 2001 to 2006 and weigh observations to give each firm an equal weight.

Figure 5: Distribution of Product × Sourcing Strategy Specific Coefficients on Sales

We display separately the distribution of all estimated coefficients (solid line) and the distribution of "precise" coefficients, which we define as the 25% coefficients with the lowest standard error (dashed line). These two distributions are uni-modal with a strictly positive mode: 0.126 and 0.114, respectively. The corresponding medians are also close to the point estimates reported in Table 6: 0.111 in the entire sample

---

[39]While we estimate one β_{σ_k} per sourcing strategy, we restrict the product-year fixed effects α_{tk} and μ_k to be common across all sourcing strategies of a same product. This increases our statistical power. The results are similar when we estimate specification (16) for each product × sourcing strategy, but the number of identifiable coefficients drops significantly.

and 0.098 in the sample of precisely estimated parameters. Moreover, about 70% of the significant coefficients are positive.[40] Hence, while many of the coefficients are imprecisely estimated given the small number of firms in the sourcing-strategy specific bins, the results are broadly consistent with our findings above: on average larger firms bias their spending toward their top variety.

To sum up, the results in Tables 5 and 6 as well as Figure 5 show that the homotheticity property of standard firm-based models of importing is systematically rejected. Not only does the data show ample dispersion in firms' spending pattern across a common set of trading partners, but these differences across firms are partly systematic. In particular, firm size, as measured by sales, has a significant effect on the distribution of expenditure across foreign varieties after controlling for the sourcing strategy. Moreover, our results suggest a particular direction for the effect of firm size on expenditure shares: in all specifications we find that larger firms spend relatively more on their top varieties. As expenditure shares are increasing in price-adjusted input quality, these results suggest that large firms bias their expenditure towards varieties of high quality.

4   Towards A Mechanism: Quality-Productivity Complementarity

The findings of Section 3 are inconsistent with any theory in the class of Section 2 where input prices are common across firms and productivity is factor neutral. They suggest that price-adjusted input qualities have an important firm-component and that their distribution systematically varies with firm size, holding the sourcing set constant. In this last section, we provide additional evidence that these properties are also present in firms' unit values. We show that, holding the sourcing strategy fixed, larger firms pay higher prices for their foreign inputs and that this effect is stronger for their most important varieties. To the extent that prices reflect quality, our findings are consistent with a complementarity between firm productivity and input quality as for example stressed by Kugler and Verhoogen (2012).[41]

We start by investigating whether larger firms pay higher prices for their inputs. We are only able to consider foreign inputs because our data does not contain information on the unit values of firms' domestic inputs. We consider the specification

$$\ln\left(p_{ickt}\right) = \alpha_S + \alpha_t + \alpha_k + \alpha_c + \alpha_\Sigma + \beta \times \ln\left(S_{it}\right) + \vartheta \times \ln\left(\text{Shipment size}_i\right) + X'_{it}\mu + u_{ickt}, \qquad (19)$$

where p_{ickt} is the price firm i pays for product k inputs from country c in year t and $\ln\left(S_{it}\right)$ denotes log-sales as previously. Note that (19) also contains a county fixed effect α_c, i.e. we compare variety-specific prices paid by firms of different size. We also control for the size of the shipment to ensure that bigger firms do not pay different prices simply because they order in bulk. 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.[42]

---

[40]For the entire sample, 23% of the estimated β_{σ_k} (representing 24% of total import value in this same sample) are significantly different from zero and 16% (i.e. 70% of the significant coefficients, representing 18% of total import value) are significantly positive, while 7% (representing 7% of import value) are significantly negative. For the "precisely estimated" sample, 27% of all coefficients (representing 21% of import value in this sample) are significant and 19% are significantly positive (representing 17% of import value) while 7% are significantly negative (representing 5% of import value).

[41]Variation in prices has been linked to variation in product quality. See Schott (2004), Hummels and Klenow (2005), Hallak and Schott (2011) or Khandelwal (2010) for evidence at the country level and Manova and Zhang (2012) or Kugler and Verhoogen (2012) for evidence at the firm level.

[42]Kugler and Verhoogen (2012) proxy productivity by employment while we use sales in our main specification, for consistency with section 3 and because unit values and sales are sourced from two different datasets. This mitigates the risk of correlation in measurement error across both variables, as in Kugler and Verhoogen (2012). We however provide the results obtained with either employment or TFP as alternative proxies for firm efficiency in Appendix C.4 and obtain qualitatively similar results, but

In contrast to them, we focus on firms' *international* input purchases. This allows us to control for firms' extensive margin, Σ, and hence for differences in firms' demand for quality induced by other inputs in the production function.[43]

The results are contained in Table 7. 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 coefficient in column 2 exceeds the unconditional one in column 1. The rest of the columns control for the sourcing strategy and add further controls. Overall, we estimate an elasticity of prices with respect to firm sales of about 7.5%.

Dependent variable:			$\ln p_{ickt}$		
	(1)	(2)	(3)	(4)	(5)
ln Sales	0.024***	0.073***	0.076***	0.080***	0.077***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln Shipment Quantity		-0.224***	-0.220***	-0.211***	-0.211***
		(0.001)	(0.001)	(0.001)	(0.001)
Other controls:	No	No	No	No	Yes
Fixed effects:	Years	Years	Years	Years	Years
	Industry	Industry	Industry	Industry ×	Industry ×
	Products	Products	Products ×	Products ×	Products ×
	Countries	Countries	Countries	Countries	Countries
Impact of a 1sd increase in sales (from mean):					
	0.063	0.194	0.203	0.213	0.206
Observations	2,195,456	2,195,456	2,195,456	2,195,456	2,141,864
Identified FE	10,998	10,998	188,481	620,341	611,882
of which singletons	3	3	63,517	309,873	306,652

Notes: Regressions at the importer × product × country, 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 foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). All regressions include industry fixed at the four digit level, product fixed effects at the eight digit level, year and country level fixed effects. Observations are weighted to give each firm an equal weight.

Table 7: Unit Values and Firm Size

The findings in Table 7 suggest that larger firms tend to buy expensive products from particular countries relative to their smaller counterparts. We now assess whether this effect is particularly strong for the most important varieties, as suggested by our earlier results on expenditure shares in Tables 5 and 6. That is, we evaluate whether large firms face a distribution of unit values which is more skewed than that of small firms. We consider the following specification:

$$\ln\left(\frac{p_{ikt}^{(1)}}{p_{ikt}^{(j)}}\right) = \alpha_S + \alpha_t + \alpha_k + \alpha_\Sigma + \beta \times \ln(S_{it}) + \vartheta_1 \ln\left(\text{Ship. size}_{it}^{(1)}\right) + \vartheta_j \ln\left(\text{Ship. size}_{it}^{(j)}\right) + X_{it}'\mu + u_{ikt}, \quad (20)$$

where $p_{ikt}^{(1)}$ is the unit value of the main variety and $p_{ikt}^{(j)}$ is the unit value of the variety of order j. Equation (20) is therefore a direct extension of Equations (17) and (19). We control for the sourcing strategy by including industry × product × sourcing strategy fixed effects, as in Panel B of Table 5. Table 8 contains the

somewhat lower magnitudes.

[43]Specifically, if countries differ in the quality of the inputs they provide and inputs are not perfect substitutes, the demand for quality from country c will depend on the quality of other inputs sourced. By controlling for the sourcing strategy Σ we control for the country-component of the other inputs.

results. We find that the coefficient β is positive in all specifications and that its magnitude increases with the order of the statistic in columns (2) to (5). Thus, the distribution of prices paid by large firms for their ranked varieties stochastically dominates that of smaller firms. In other words, large firms have particularly high unit values for their most important foreign varieties.

Dependent variable:	$\ln\left(p_{ikt}^{(1)}\right)$	$\ln\left(p_{ikt}^{(1)}/p_{ikt}^{(2)}\right)$	$\ln\left(p_{ikt}^{(1)}/p_{ikt}^{(3)}\right)$	$\ln\left(p_{ikt}^{(1)}/p_{ikt}^{(4)}\right)$	$\ln\left(p_{ikt}^{(1)}/p_{ikt}^{(5)}\right)$
	(1)	(2)	(3)	(4)	(5)
ln Sales	0.119***	0.050***	0.084***	0.121***	0.200***
	(0.004)	(0.004)	(0.011)	(0.021)	(0.044)
ln Shipment Quantity	-0.285***	-0.378***	-0.396***	-0.455***	-0.511***
Variety 1	(0.003)	(0.004)	(0.008)	(0.013)	(0.026)
ln Shipment Quantity	-	0.338***	0.311***	0.339***	0.359***
Variety k		(0.003)	(0.006)	(0.011)	(0.017)
Impact of a 1sd increase in sales (from mean):					
	0.316	0.132	0.212	0.292	0.465
Observations	711,764	702,310	291,300	150,788	88,080
Identified FE	516,258	509,701	248,315	137,584	83,198
of which singletons	421,628	416,332	221,666	128,216	79,443

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. 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 also include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product). Observations are weighted to give each firm an equal weight, independently of the number of imported products.

Table 8: Relative Unit Values and Firm Size

Taken together, these findings suggest that firms employ different suppliers within foreign countries. In particular, larger firms tend to connect with more expensive input suppliers in every country, which may provide products of higher quality. Furthermore, this pattern is especially strong for the most important varieties they source. These findings may be the result of a number of underlying mechanisms, such as the process of supplier search or a pattern of relationship-specific investments that firms undertake with their most important suppliers. In fact, in Section C.5 we report a decomposition of firm-level unit values and show that the cross-firm variation in input prices within countries is large relative to the cross-country variation. Distinguishing between alternative micro mechanisms of supplier choice within country is beyond the scope of this paper as it would require more information on firms' suppliers than available to us.

5 Conclusion

In this paper, we use micro-data for French manufacturing firms to study the intensive margin of firms' import demand. We find that firms' import demand system is non-homothetic: holding the extensive margin of trade constant, firm size directly affects the distribution of spending across the set of trading partners. This finding is inconsistent with virtually all firm-based models of importing that have been considered in the literature. In particular, it rejects a class of theories that assume that productivity is factor neutral, production is subject to constant returns to scale and prices are common across firms, as such models imply that expenditure shares are fully determined by price-adjusted qualities, that is, by characteristics of the supplying countries.

Moreover, the particular direction in which the theory is rejected is economically meaningful. We first show that, for a given sourcing strategy, larger firms have lower import shares. This is in stark contrast

to the behavior of import spending stemming from the extensive margin, as large firms source from more international markets. Because firms' import shares summarize the productivity gains from importing, this result implies that large firms gain less from importing holding the set of trading partners constant. Secondly, we turn to the allocation of spending across foreign input suppliers and show that the concentration of import spending is also increasing in firm size. In particular, larger firms spend a higher share of their budget on their most important supplier. Finally, we turn to the patterns of input prices. We show that the within-country price variation across importing firms is large relative to the variation in average prices across countries, that prices are positively correlated with firm size and that large firms pay particularly high prices for their most important inputs. These patterns are consistent with a complementarity between input quality and firm productivity whereby larger firms match with high quality suppliers within foreign countries.

References

Amiti, Mary and Jozef Konings, "Trade Liberalization, Intermediate Inputs and Productivity," *American Economic Review*, 2007, *97* (5), 1611–1638.

_ , 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. forthcoming.

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

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, "The Empirics of Firm Heterogeneity and International Trade," *Annual Review of Economics*, 2012, *4* (1), 283–313.

Bernard, Andrew B, J Bradford Jensen, Stephen J Redding, and Peter K Schott, "Global Firms," Technical Report, National Bureau of Economic Research 2016.

Bernard, Andrew B., J. Bradford Jensen, Stephen Redding, and Peter Schott, "Firms in International Trade," *Journal of Economic Perspectives*, 2007, *3* (21), 105–130.

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "The Gains from Input Trade with Heterogeneous Importers," CEPR Discussion Papers 11721 2016.

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," 2015. Working Paper.

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.

Hallak, Juan Carlos and Peter K. Schott, "Estimating Cross-Country Differences in Product Quality," *Quarterly Journal of Economics*, 2011, *126*, 417–474.

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

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

Hummels, David and Pete Klenow, "The Variety and Quality of a Nation's Exports," *The American Economic Review*, 2005, *95* (3), 704–723.

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," *Review of Economic Studies*, 2010, *77*, 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.

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

Pierce, Justin R. and Peter K. Schott, "Concording U.S. Harmonized System Codes over Time," *Journal of Official Statistics*, 2012, *28* (1), 53–68.

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

Schott, Peter, "Across-Product Versus Within-Product Specialization in International Trade," *Quarterly Journal of Economics*, 2004, *119* (2), 647–678.

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

Appendix

A 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

$$\frac{1}{\xi_{ck}} = \lambda \varphi \frac{\partial \hat{f}(x)}{\partial x_k} \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{ck}}, \tag{21}$$

where λ is the multiplier on the constraint. Consider a product k. Then

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

where the last equality follows from constant returns of $\hat{g}_k$.[44] Summing across products,

$$\Gamma(\Sigma, x, \varphi) = \sum_{\{c \in \Sigma_k\}_k} \tilde{z}_{ck}/\xi_{ck} = \lambda \varphi \sum_{k=1}^{n} \frac{\partial \hat{f}(x)}{\partial x_k} x_k = \lambda \varphi \hat{f}(x) = \lambda x, \tag{22}$$

so that λ 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 $[\xi_{ck}]$. (21) implies that

$$\left[\lambda \varphi \frac{\partial \hat{f}(x)}{\partial x_k} \right]^{-1} = \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{c'k}} \xi_{c'k} = \frac{\partial \hat{g}_k(\tilde{z})}{\partial \tilde{z}_{ck}} \xi_{ck} \text{ for all } c, c' \in \Sigma_k. \tag{23}$$

As $\hat{g}_k$ is constant returns to scale, (23) implies that

$$\frac{\partial \hat{g}_k(1, \chi_{2k}, .., \chi_{C_k k})}{\partial \tilde{z}_{ck}} = \frac{\partial \hat{g}_k(1, \chi_{2k}, .., \chi_{C_k k})}{\partial \tilde{z}_{1k}} \frac{\xi_{1k}}{\xi_{ck}} \text{ for all } c \in \Sigma_k. \tag{24}$$

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

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

---

[44]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 μ yields

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

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

B   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.[45] 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 Simplifé 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 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 investments in other firms (participation) are higher than 1.2 million euros or have sales above 60 million euros or have more 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 × 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 lest one product.

---

[45]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.

C   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.

C.1   The Variety Margin

Import Trade Interactions.   Table 9 complements Table 1 and contains descriptive statistics on trade flows at the variety level.[46]   Table 9 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 × product cells. Similarly, for half of the potential products (i.e. roughly 5,000 products) only 28 country × firm interactions are observed, while the most popular products are shipped into France in more than 800 distinct country-firm combinations.[47]

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 × 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.

			Quantiles			
Variable	Dimension	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 in the Appendix for a complete description of the data.

Table 9: The Concentration of French Imports at the Variety Level

Distribution of Spending.   How do firms allocate their expenditure (for one considered product) across source countries, i.e. across varieties? Figure 6 complements the two last rows of Table 9 and depicts the 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.

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

---

[46]The statistics in Table 9 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.

[47]Note in particular the still very large difference between the 99th and 90th quantiles.
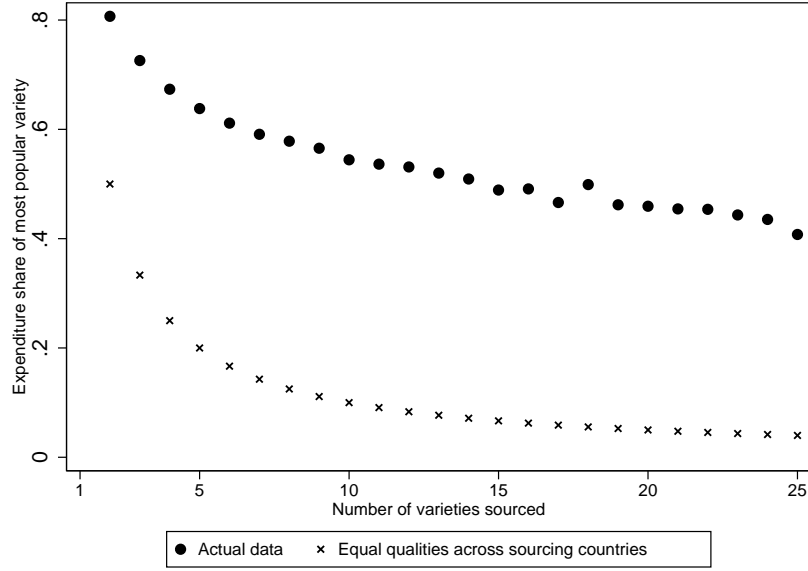
Notes: The figure shows the average expenditure share on the top variety (y-axis) of firm×product pairs with varying number of varieties (x-axis). We also depict the counterfactual expenditure share if price-adjusted quality flows (and hence expenditure shares) were equalized across varieties. We use 6 years of data from 2001-2006 and report the yearly average.

Figure 6: Concentration of Firms' Import Spending on Main Variety Across Sourcing Strategies

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.

C.2 Robustness Checks and Extensions for the Analysis of the Intensive Margin of Imports

Firm Fixed Effects. Table 10 first replicates the baseline analysis of Table 5 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.[48] Reassuringly, the obtained results are very similar to those of the baseline specification, in spite of the different identification strategy.

Panel (A): Controlling for the Number of Countries

	$\ln\left(\frac{s_1}{1-s_1}\right)$	$\ln\left(\frac{s_1}{s_2}\right)$	$\ln\left(\frac{s_1}{s_3}\right)$	$\ln\left(\frac{s_1}{s_4}\right)$	$\ln\left(\frac{s_1}{s_5}\right)$
	(1)	(2)	(3)	(4)	(5)
ln Sales	0.130***	0.128***	0.103***	0.120***	0.075
	(0.015)	(0.016)	(0.030)	(0.040)	(0.055)
ln Nb countries	-1.346***	-0.792***	-1.560***	-2.046***	-2.543***
	(0.014)	(0.014)	(0.029)	(0.045)	(0.064)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.348	0.341	0.262	0.292	0.175
- On s_k	0.058	-0.048	-0.015	-0.010	-0.003
Sample average of s_k	0.756	0.206	0.080	0.041	0.024
Observations	739,167	739,167	306,573	158,652	92,722
Identified FE	481,164	481,164	218,488	116,965	69,242
of which singletons	343,330	343,330	161,455	87,971	52,552

Panel (B): Controlling for the Entire Sourcing Strategy

	$\ln\left(\frac{s_1}{1-s_1}\right)$	$\ln\left(\frac{s_1}{s_2}\right)$	$\ln\left(\frac{s_1}{s_3}\right)$	$\ln\left(\frac{s_1}{s_4}\right)$	$\ln\left(\frac{s_1}{s_5}\right)$
	(1)	(2)	(3)	(4)	(5)
ln Sales	0.110***	0.109***	0.110**	0.313***	0.085
	(0.019)	(0.019)	(0.048)	(0.071)	(0.122)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.294	0.290	0.278	0.760	0.198
- On s_k	0.050	-0.042	-0.016	-0.020	-0.002
Sample average of s_k	0.756	0.206	0.080	0.041	0.024
Observations	739,167	739,167	306,573	158,652	92,722
Identified FE	552,033	552,033	267,749	146,852	88,265
of which singletons	439,840	439,840	235,963	136,334	84,276

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. 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). Regressions in panel (A) include fixed effects for years, 4-digit industries interacted with 8-digit products and sourcing countries (of the specific varieties considered), as well as firm fixed effects. Regressions in panel (B) include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product), as well as firm fixed effects. 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 10: Firm Size and the Intensive Margin of Trade, Controlling for Firm Fixed Effects

---

[48]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.

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 9 and Figure 6. In Table 11 however, we present the results obtained for our central specification (Table 5) without weighting. The obtained results are almost unaffected, both qualitatively and quantitatively.

Panel (A): Controlling for the Number of Countries

	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
ln Sales	0.106***	0.099***	0.133***	0.154***	0.182***
	(0.004)	(0.004)	(0.007)	(0.011)	(0.016)
ln Nb countries	-1.388***	-0.857***	-1.582***	-2.050***	-2.503***
	(0.009)	(0.009)	(0.018)	(0.029)	(0.042)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.230	0.215	0.271	0.298	0.337
- On s_k	0.040	-0.030	-0.016	-0.009	-0.006
Sample average of s_k	0.753	0.200	0.079	0.041	0.024
Observations	739,167	739,167	306,573	158,652	92,722
Identified FE	462,159	462,159	208,039	111,284	66,048
of which singletons	339,672	339,672	158,378	85,786	51,169

Panel (B): Controlling for the Entire Sourcing Strategy

	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
ln Sales	0.127***	0.125***	0.174***	0.185***	0.165***
	(0.005)	(0.005)	(0.014)	(0.026)	(0.051)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.277	0.272	0.354	0.358	0.305
- On s_k	0.048	-0.038	-0.021	-0.011	-0.006
Sample average of s_k	0.753	0.200	0.079	0.041	0.024
Observations	739,167	739,167	306,573	158,652	92,722
Identified FE	534,061	534,061	261,019	144,718	87,588
of which singletons	435,837	435,837	232,96	134,865	83,661

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. 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). Regressions in panel (A) include fixed effects for years, 4-digit industries interacted with 8-digit products and sourcing countries (of the specific varieties considered). Regressions in panel (B) include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product). 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 11: Firm Size and the Intensive Margin of Trade, Unweighted Regressions

Alternative Proxies for Productivity. In our central specification of Table 5, firm sales are first and foremost a proxy for overall firm size, a firm level characteristic which enable to test Proposition 1. In this Appendix, we investigate whether our baseline results are robust to other proxies for firm productivity. In Table 12 we use employment as an alternative measure of firm size.[49] In Table 13 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.

Panel (A): Controlling for the Number of Countries

	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
ln Employment	0.099***	0.092***	0.142***	0.171***	0.182***
	(0.006)	(0.006)	(0.011)	(0.017)	(0.023)
ln Nb countries	-1.441***	-0.874***	-1.657***	-2.126***	-2.662***
	(0.015)	(0.015)	(0.027)	(0.043)	(0.059)
Impact of a 1 standard deviation increase in employment (from mean):					
- On dependent variable	0.189	0.175	0.258	0.300	0.311
- On s_k	0.033	-0.026	-0.016	-0.010	-0.006
Sample average of s_k	0.756	0.206	0.080	0.041	0.024
Observations	737,248	737,248	305,828	158,261	92,467
Identified FE	460,937	460,937	207,544	111,032	65,881
of which singletons	338,698	338,698	157,989	85,600	51,04

Panel (B): Controlling for the Entire Sourcing Strategy

	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
ln Employment	0.104***	0.102***	0.154***	0.151***	0.163**
	(0.008)	(0.008)	(0.020)	(0.040)	(0.066)
Impact of a 1 standard deviation increase in employment (from mean):					
- On dependent variable	0.198	0.195	0.280	0.266	0.278
- On s_k	0.035	-0.029	-0.017	-0.009	-0.006
Sample average of s_k	0.756	0.206	0.080	0.041	0.024
Observations	737,248	737,248	305,828	158,261	92,467
Identified FE	532,655	532,655	260,365	144,351	87,340
of which singletons	434,628	434,628	232,355	134,518	83,420

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. 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, (ln) intermediate inputs per worker and an indicator of capital intensity (ln tangible assets per worker). Regressions in panel (A) include fixed effects for years, 4-digit industries interacted with 8-digit products and sourcing countries (of the specific varieties considered). Regressions in panel (B) include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product). Observations are weighted to give each firm an equal weight, independently of the number of imported products. TFP is estimated at the industry 2 digit level using the Petrin and Levinsohn (2012) methodology, on the sample of all French manufacturing firms (both importing and non-importing). 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 12: Firm Size and the Intensive Margin of Trade, Using Employment as a Proxy for Firm Size

---

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

Panel (A): Controlling for the Number of Countries

	$\ln\left(\frac{s_1}{1-s_1}\right)$	$\ln\left(\frac{s_1}{s_2}\right)$	$\ln\left(\frac{s_1}{s_3}\right)$	$\ln\left(\frac{s_1}{s_4}\right)$	$\ln\left(\frac{s_1}{s_5}\right)$
	(1)	(2)	(3)	(4)	(5)
ln TFP	0.231***	0.222***	0.328***	0.491***	0.547***
	(0.031)	(0.031)	(0.059)	(0.087)	(0.125)
ln Nb countries	-1.385***	-0.823***	-1.562***	-1.985***	-2.477***
	(0.017)	(0.017)	(0.031)	(0.048)	(0.065)
Impact of a 1 standard deviation increase in TFP (from mean):					
- On dependent variable	0.107	0.103	0.169	0.195	0.221
- On s_k	0.019	-0.015	-0.010	-0.006	-0.004
Sample average of s_k	0.755	0.206	0.080	0.041	0.024
Observations	588,820	588,820	244,644	126,947	74,335
Identified FE	381,856	381,856	171,227	91,638	54,404
of which singletons	283,730	283,73	131,685	71,399	42,552

Panel (B): Controlling for the Entire Sourcing Strategy

	$\ln\left(\frac{s_1}{1-s_1}\right)$	$\ln\left(\frac{s_1}{s_2}\right)$	$\ln\left(\frac{s_1}{s_3}\right)$	$\ln\left(\frac{s_1}{s_4}\right)$	$\ln\left(\frac{s_1}{s_5}\right)$
	(1)	(2)	(3)	(4)	(5)
ln TFP	0.236***	0.234***	0.285***	0.588***	0.583*
	(0.038)	(0.038)	(0.099)	(0.173)	(0.336)
Impact of a 1 standard deviation increase in TFP (from mean):					
- On dependent variable	0.110	0.109	0.147	0.234	0.236
- On s_k	0.020	-0.016	-0.009	-0.007	-0.004
Sample average of s_k	0.755	0.206	0.080	0.041	0.024
Observations	588,820	588,820	244,644	126,947	74,335
Identified FE	435,964	435,964	210,613	116,501	70,404
of which singletons	358,144	358,144	188,657	108,847	67,312

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. 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). Regressions in panel (A) include fixed effects for years, 4-digit industries interacted with 8-digit products and sourcing countries (of the specific varieties considered). Regressions in panel (B) include fixed effects for years, four-digit industries interacted with 8-digit products and entire sourcing strategies (for the product). 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: Firm TFP and the Intensive Margin of Trade

C.3 CES Approach

One challenge of the statistical analysis of the intensive margin of imports in Section 3.4 is to control as accurately as possible for the extensive margin. In the baseline regressions of Table 5, we insert high-dimensional sets of fixed effects, while the approach underlying Table 6 focuses even more precisely on firms' intensive margin problem by considering only firms × 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 14 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:

$$\ln \left(\frac{s^i_{jk}}{s^i_{j'k}} \right) = \ln \left(s^i_{jk} \right) - \ln \left(s^i_{j'k} \right) = (\rho - 1) \left(\ln \left(\xi_{jk} \right) - \ln \left(\xi_{j'k} \right) \right), \tag{25}$$

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.[50] While equation (25) requires stronger assumptions than the general expenditure share equation (8) of Section 2, it is very useful. Indeed, the key advantage of (25) over (17) 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.

   To make (25) 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 6. We then rank these V countries for each firm and run the regression contained in (17) on the enlarged subsamples.[51] Table 14, which has the same structure as Table 6, 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 6, we also recover the monotonicity of the coefficients and the point estimates are very similar in magnitude. Hence, in Table 14 again, larger firms appear to concentrate their spending on their most preferred sourcing countries relative to smaller importers.

   We also exploit the CES structure to extend the analysis in the main text which obtains firm-size coefficients that vary by sourcing strategy. That is, for each product and pair of varieties, we can pool all firms whose sourcing strategies include at least such pair of varieties. We estimate the following equation for each

---

[50]Of course (25) is only defined for pairs of varieties which are actually sourced. That is, expression (25) is valid for $c, c' \in \Sigma_{ki}$, and not defined otherwise. This means that when testing for the exclusion restriction implicit in (25) 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.

[51]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.

|	$\ln\left(\frac{s_1}{1-s_1}\right)$ (1)	$\ln\left(\frac{s_1}{s_2}\right)$ (2)	$\ln\left(\frac{s_1}{s_3}\right)$ (3)	$\ln\left(\frac{s_1}{s_4}\right)$ (4)	$\ln\left(\frac{s_1}{s_5}\right)$ (5)
Firms sourcing (at least) from the 2 most popular countries					
ln Sales	0.089***	0.089***			
	(0.005)	(0.005)			
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.225	0.225			
- On s_k	0.043	-0.032			
Sample average of s_k	0.716	0.208			
Observations	210,341	210,341			
Identified FE	10,765	10,765			
of which singletons	1,371	1,371			
Firms sourcing (at least) from the 3 most popular countries					
ln Sales	0.051***	0.041***	0.127***		
	(0.006)	(0.006)	(0.009)		
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.120	0.098	0.299		
- On s_k	0.027	-0.012	-0.021		
Sample average of s_k	0.633	0.218	0.091		
Observations	79,453	79,453	79,453		
Identified FE	9,180	9,180	9,180		
of which singletons	2,107	2,107	2,107		
Firms sourcing (at least) from the 4 most popular countries					
ln Sales	0.027***	0.014*	0.058***	0.142***	
	(0.008)	(0.008)	(0.010)	(0.013)	
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.060	0.031	0.130	0.319	
- On s_k	0.014	-0.001	-0.010	-0.013	
Sample average of s_k	0.588	0.217	0.102	0.050	
Observations	40,614	40,614	40,614	40,614	
Identified FE	7,446	7,446	7,446	7,446	
of which singletons	2,299	2,299	2,299	2,299	
Firms sourcing (at least) from the 5 most popular countries					
ln Sales	0.021**	0.009	0.026*	0.078***	0.151***
	(0.010)	(0.010)	(0.013)	(0.016)	(0.020)
Impact of a 1 standard deviation increase in sales (from mean):					
- On dependent variable	0.044	0.019	0.056	0.169	0.326
- On s_k	0.011	0.000	-0.004	-0.008	-0.008
Sample average of s_k	0.560	0.212	0.107	0.058	0.031
Observations	23,989	23,989	23,989	23,989	23,989
Identified FE	5,967	5,967	5,967	5,967	5,967
of which singletons	2,219	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 14: Firm Size and the Intensive Margin of Trade, Exploiting the CES Structure

product × couple of countries:

$$\ln\left(\frac{s^i_{1k,\{c,c'\}}}{1 - s^i_{1k,\{c,c'\}}}\right) = \alpha_{k,\{c,c'\}} + \beta_{k,\{c,c'\}}\ln S_{it} + X'_{it}\mu_{k,\{c,c'\}} + u_{ik,\{c,c'\}}, \text{ for all } k \text{ and } c \neq c' \qquad (26)$$

This leaves us with 54,755 coefficients to be estimated, which are represented on Figure 7. As in Figure 5, we display separately the distribution of all estimated coefficients (solid line) and the distribution of "precise" coefficients, which we define as the 25% coefficients with the lowest standard error (dashed line). These two distributions are uni-modal with a strictly positive mode: 0.128 and 0.125, respectively. The corresponding medians are also close to the point estimates reported in Table 6: 0.129 in the entire sample and 0.101 in the sample of precisely estimated parameters. Moreover, about 60% of the significant coefficients are positive.[52]
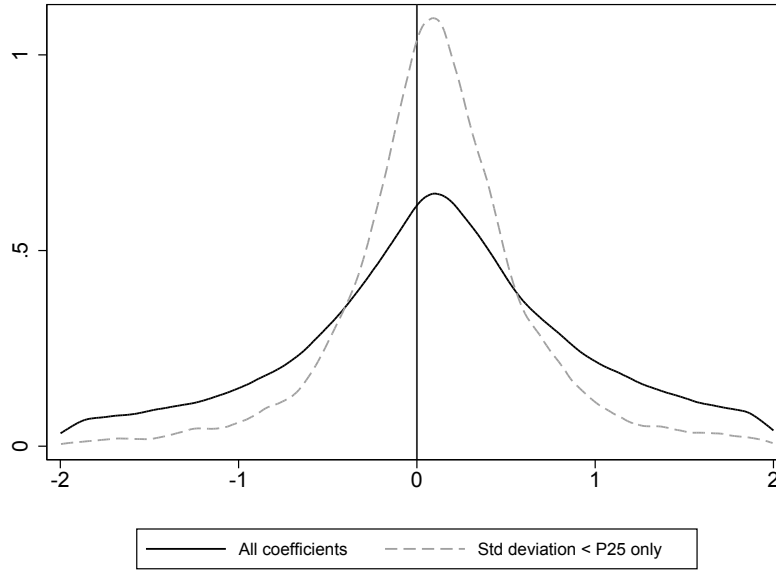
Notes: This figure represents the distribution of product × {c,c'} coefficients obtained from the regression (26). Only sourcing strategies × products with more than 10 observations are included into the sample. We consider the years 2001 to 2006 and weigh observations to give each firm an equal weight.

Figure 7: Distribution of Product × $\{c, c'\}$ Specific Coefficients on Sales in CES Specifications

---

[52]For the entire sample, 31% of the estimated $\beta_{k,\{c,c'\}}$, representing 33% of import expenditures, are significantly different from zero and 19% (representing 21% of import expenditures, and 61% of the number of significant coefficients) are significantly positive, while 12% are significantly negative (representing 11% of import expenditures). For the "precisely estimated" sample, 42% of all coefficients are significant and 27% are significantly positive while 13% are significantly negative.

C.4 Unit Values and Firm Size

Specifications with Firm Fixed Effects. Table 15 first proposes an extension of the analysis in Section 4 where we insert an additional set of firm fixed effects in Equation (19). This implies that in these specifications, the coefficients of interest are identified from *variation* in both unit values and either firm sales or shipment quantities. The results obtained in these alternative specifications are almost unaffected.

Dependent variable:			$\ln p_{ickt}$		
	(1)	(2)	(3)	(4)	(5)
ln Sales	0.016***	0.060***	0.068***	0.074***	0.091***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
ln Shipment Quantity		-0.205***	-0.206***	-0.198***	-0.198***
		(0.001)	(0.001)	(0.001)	(0.001)
Other controls:	No	No	No	No	Yes
Fixed effects:	Years	Years	Years	Years	Years
	Industry	Industry	Industry	Industry ×	Industry ×
	Products	Products	Products ×	Products ×	Products ×
	Countries	Countries	Countries	Countries	Countries
Impact of a 1sd increase in sales (from mean):					
	0.041	0.159	0.181	0.198	0.242
Observations	2,195,456	2,195,456	2,195,456	2,195,456	2,141,864
Identified FE	38,966	38,966	216,348	646,114	637,030
of which singletons	4	4	63,907	311,112	307,837

Notes: Regressions at the importer × product × country, 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 foreign headquarters, and an indicator of capital intensity (tangible assets per worker). All regressions include industry fixed at the four digit level, product fixed effects at the eight digit level, year and country level fixed effects as well as firm fixed effects. Observations are weighted to give each firm an equal weight.

Table 15: Unit Values and Firm Size, Controlling for Firm Fixed Effects

Specifications with Alternative Proxies for Firm Size or Productivity. In Tables 16 and 17, we present the estimation results obtained for Equation (19) with an alternative proxy of firm size (employment) or productivity (TFP à la Petrin and Levinsohn, 2012). Again, the regression results show that our findings are largely preserved qualitatively, although the quantitative implications are slightly lower.[53]

---

[53]Notice that since the information about sales is not taken from the same data set as the information about import spending, the difference in the results is unlikely to be mainly driven by correlation in measurement errors across both sets of variables, as suggested in Kugler and Verhoogen (2012).

Dependent variable:			$\ln p_{ickt}$		
	(1)	(2)	(3)	(4)	(5)
ln Employment	0.036***	0.067***	0.070***	0.071***	0.069***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ln Shipment Quantity		-0.222***	-0.218***	-0.208***	-0.210***
		(0.001)	(0.001)	(0.001)	(0.001)
Other controls:	No	No	No	No	Yes
Fixed effects:	Years	Years	Years	Years	Years
	Industry	Industry	Industry	Industry ×	Industry ×
	Products	Products	Products ×	Products ×	Products ×
	Countries	Countries	Countries	Countries	Countries
Impact of a 1sd increase in employment (from mean):					
	0.068	0.127	0.132	0.135	0.131
Observations	2,184,904	2,184,904	2,184,904	2,184,904	2,136,058
Identified FE	10,995	10,995	188,058	618,156	610,732
of which singletons	3	3	63,362	308,901	306,077

Notes: Regressions at the importer × product × country, 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 foreign headquarters,(ln) intermediate inputs per workers and an indicator of capital intensity (ln tangible assets per worker). All regressions include industry fixed at the four digit level, product fixed effects at the eight digit level, year and country level fixed effects as well as firm fixed effects. Observations are weighted to give each firm an equal weight.

Table 16: Unit Values and Firm Size, Using Employment as a Proxy for Firm Size

Dependent variable:			$\ln p_{ickt}$		
	(1)	(2)	(3)	(4)	(5)
ln TFP	0.068***	0.131***	0.123***	0.108***	0.098***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
ln Shipment Quantity		-0.220***	-0.215***	-0.204***	-0.206***
		(0.001)	(0.001)	(0.001)	(0.001)
Other controls:	No	No	No	No	Yes
Fixed effects:	Years	Years	Years	Years	Years
	Industry	Industry	Industry	Industry ×	Industry ×
	Products	Products	Products ×	Products ×	Products ×
	Countries	Countries	Countries	Countries	Countries
Impact of a 1sd increase in TFP (from mean):					
	0.060	0.114	0.107	0.095	0.086
Observations	1,748,936	1,748,936	1,748,936	1,748,936	1,748,936
Identified FE	10,302	10,302	167,540	536,776	536,776
of which singletons	4	4	56,435	270,171	270,171

Notes: Regressions at the importer × product × country, 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 foreign headquarters, and an indicator of capital intensity (ln tangible assets per worker). All regressions include industry fixed at the four digit level, product fixed effects at the eight digit level, year and country level fixed effects as well as firm fixed effects. TFP is estimated at the industry 2 digit level using the Petrin and Levinsohn (2012) methodology, on the sample of all French manufacturing firms (both importing and non-importing). Observations are weighted to give each firm an equal weight.

Table 17: Unit Values and Firm Productivity

C.5 The Variance of Foreign Prices

In Table 18, we propose a decomposition of the variance of unit values. We proceed in the same way as we did for firms' domestic shares in Table 2 or firms' foreign expenditure Shares in Table 4. In column 1 we show that, unsurprisingly, a large part of the variation in prices is across products.[54] In particular, product-level

---

[54]This is unsurprising since these fixed effects capture differences in measurement units across products in the index of quantity of the custom data. Therefore, this amount of heterogeneity is hardly interpretable.

fixed effects capture almost three quarters of the variation. More surprisingly, column 2 shows that country fixed effect contribute little even though they presumably capture differences in transportation costs across different varieties of the same product. When inserted on top of the product level fixed effects (column 2), they explain 1.3 percentage point of the variance, while more than 25% of the variance of prices remains un-explained. When we control for an interaction of country and product level fixed effects, they explain 6.5 percentage point of the variance (column 3), while again, 20% remains un-explained. Therefore, average prices appear to be relatively homogeneous across countries (Waugh, 2010; Hsieh and Klenow, 2007), or rather the within-country variation is large relative to the cross-country variation (once controlled for product level fixed effects).

In columns 4 to 7 we investigate the nature of this large residual within countries. In particular, we show that there is an important firm-component driving this variation. Firm-level fixed effects alone explain about 5% of the variation conditional on a full set of product-country interactions (column 4). Their explanatory power increases when interacted with product and/or country level effects (columns 5 - 7). These patterns of firm-specific prices within country-product pairs are consistent with quantitatively meaningful firm-level price biases (which are correlated with firm size) as documented in Tables 7 and 8. They are also consistent with firm-level disagreement in the ranking of varieties that is documented in Section 3.

				$\ln p_{ickt}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Observations	2,195,456	2,195,456	2,195,456	2,195,456	2,195,456	2,195,456	2,195,456
Identified FE	10,428	10,674	18,8157	216,324	337,688	259,400	1,167,534
of which singletons	2	3	63,517	63,907	4	78,416	698,564
Variance decomposition, removing singletons (Correia, 2015; Gaure, 2013):							
R^2	0.730	0.743	0,795	0.849	0.880	0.868	0.942
Fixed effects 1:	0.730	0.705	0.000	0.107	0.861	0.194	0.000
Fixed effects 2:	0.000	0.015	0.795	0.581	0.005	0.476	0.942
Covariance term	0.000	0.023	0.000	0.161	0.014	0.199	0.000
Fixed effects 1:	Year	Year	Year	Year	Year	Year	Year
	Product	Product		Firm	Firm ×	Firm ×	Firm ×
					Product	Country	Product ×
							Country
Fixed effects 2:	-	Country	Product ×	Product ×	Country	Product	-
			Country	Country			

Notes: Regressions at the importer × product × country, importers of at least 2 varieties of the considered product. These importers are observed from 2001 to 2006. 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 18: Variance Decomposition for Foreign Variety Prices