Non-Homothetic Import Demand: Firm Productivity and Quality Bias*

Joaquin Blaum†, Claire Lelarge‡, Michael Peters§

June 20, 2013

Abstract

We use French microdata to learn about the behavior of importers. We look at the data through the lens of a model of importing that features complementarities across imported inputs, variety-specific fixed costs and factor neutral productivity. The model is general enough to nest most of the existing frameworks of importing. We show that, in contrast to models of exporting, the theory has no general predictions for the extensive margin of importing. In particular, if foreign inputs are imperfectly substitutable, more productive firms do not necessarily enter more foreign markets than their less productive counterparts. In contrast, the intensive margin of importing features a robust testable prediction: holding the extensive margin fixed, expenditure shares across products and varieties are fully determined by price-adjusted qualities, that is, by characteristics of the supplying country. In particular, firm productivity should not affect relative input demand once the extensive margin is controlled for. Our main contribution lies in providing a careful test of this homotheticity prediction. We show that this property is not supported by the data. The particular direction in which the theory is rejected is economically meaningful: holding the extensive margin fixed, larger firms spend relatively more on their most important variety. We provide three mechanisms how available models of importing could be made consistent with this fact: a complementarity between input quality and firm productivity, a search process by which larger firms search for foreign suppliers more intensively, and the presence of intra-firm trade.

*We especially thank Arnaud Costinot for numerous helpful discussions and persistent encouragement. We also thank Daron Acemoglu, Costas Arkolakis, Lorenzo Caliendo, Pablo Fajgelbaum, Penny Goldberg, Oleg Itskhoki, Sam Kortum, Peter Schott and seminar participants at the University of Chicago, MIT, Yale and Brown.

†Brown University. eMail: joaquin_blaum@brown.edu

‡Centre de Recherche en Économie et Statistique (CREST). eMail: claire.lelarge@ensae.fr

§Yale University and London School of Economics. eMail: m.peters@yale.edu

1 Introduction

Understanding the behavior of individual firms has been at the center of the recent research agenda in international trade. Most papers have focused on the role of firms as exporters, providing new empirical regularities and building theoretical frameworks to account for them.[1] Compared to this literature on exports, the literature on firms' import behavior is rather underdeveloped. In particular, a theory that is quantitatively consistent with the micro evidence on importing is lacking. While recent contributions have developed models of import behavior, their focus has not been to test these models but to use them for particular applications.[2] In this paper, we argue that these contributions share a common theoretical framework and we use French micro data to test its underlying assumptions. In this sense, we take a first step towards identifying the ingredients that any model of importing should have to be consistent with the micro data.

To do so, we consider a standard framework of import behavior which can speak to the firm-level data at the product and country level. Import decisions are modeled as the solution to a static maximization problem where both complementarities and fixed costs are present. Firms have a constant returns to scale production function, are price takers and have access to a set of potential inputs ("products"), each of which can be sourced domestically or from multiple sourcing countries. We refer to a particular product that is sourced from different countries as different varieties.[3] Firms' demand for imported inputs stems from quality-differences (that is, foreign varieties can be of higher quality than the domestic one) and from love-for-variety effects. In the cross-section, we allow for unrestricted firm heterogeneity in two dimensions: factor-neutral productivity and fixed costs of international sourcing.

We use this framework to study the behavior of firms as importers both at the extensive and intensive margin. We start by showing that, if the fixed costs of international sourcing are country-specific, the theory features no general predictions at the extensive margin. In particular, the extensive margin of importing is not characterized by hierarchical sourcing, i.e. less productive firms do not necessarily source their inputs from a subset of the supplying countries of their more productive counterparts. The reason is that the decision to enter a particular import market depends on the entry decisions in all other markets. This interdependency of entry across markets stems from the complementarity that inputs have in the production process. Intuitively, we can have a situation where high productivity firms source few cheap but high-fixed cost varieties, while low productivity firms source many low-fixed cost varieties, whose unit price is high. An additional consequence of

[1] See Bernard, Jensen, Redding, and Schott (2012) for a survey of the empirical literature and Redding (2011) for a summary of the theoretical literature. Recently, it has also been shown that an augmented version of the Melitz model is quantitatively consistent with several aspects of the micro data (Eaton, Kortum, and Kramarz (2011), Arkolakis (2010)).

[2] They, for example, focus on the role of imported inputs for productivity gains of trade liberalizations (Halpern, Koren, and Szeidl (2009), Amiti and Konings (2007), and Goldberg, Khandelwal, Pavcnik, and Topalova (2010)), the aggregate effects of large devaluations (Gopinath and Neiman (2012)), and the exchange rate disconnect (Amiti, Itskhoki, and Konings (2012)). We will review the related literature in more detail below.

[3] This definition is standard in the literature (see for example Goldberg, Khandelwal, Pavcnik, and Topalova (2010) and Broda and Weinstein (2006)). In the data we measure products at the 8 digit level.

this interdependency is that the firm's extensive margin problem is computationally intense.[4] These features are in stark contrast to the export problem, where entry decisions are made "market by market" and more productive firms export to a greater number of markets - see for example Eaton, Kortum, and Kramarz (2011); Arkolakis and Muendler (2011); Tintelnot (2012) and Mayer, Melitz, and Ottaviano (2010).

In contrast to the extensive margin, the intensive margin is straightforward to solve for and features sharp testable predictions under general conditions. We show that conditional on the sourcing strategy, that is the set of varieties sourced, firms' expenditure shares across products and varieties are fully determined by price-adjusted qualities, that is, by characteristics of the supplying country. Hence, the import demand system is homothetic, holding firms' extensive margin of trade fixed.[5] This property implies a strong prediction in the form of an exclusion restriction, which can be tested using firm-level data: firm productivity should not affect relative input demand once the sourcing strategy is controlled for. This exclusion restriction is a general property of any model of import behavior satisfying (i) factor neutral productivity differences and constant returns to scale in production and (ii) common input prices across firms.[6] In particular, this prediction holds regardless of how the extensive margin of trade is determined. Hence, even if firms' trading partners were the result of a dynamic process of network formation (Chaney, 2013) or matching market (Grossman, Helpman, and Kircher, 2013), the theory would still be characterized by this exclusion restriction, as long as assumptions (i) and (ii) are satisfied. Note also that we do not need to impose any assumptions on the nature of competition in product markets or on the distribution of fixed costs and input prices and qualities. This implication on the intensive margin is therefore contained in virtually all of the existing contributions of the importing literature, e.g. Goldberg, Khandelwal, Pavcnik, and Topalova (2010), Halpern, Koren, and Szeidl (2009) and Gopinath and Neiman (2012).

The main contribution of this paper is to provide a careful test of this exclusion restriction. Implementing this test, however, is not entirely straightforward. The difficulty lies in appropriately controlling for the sourcing strategy, that is, for the particular set of varieties that are sourced by the firm. We consider three complementary approaches to tackle this issue. First, we proxy the sourcing strategy by its cardinality, that is, the number of varieties sourced. While the number of sourcing countries is indeed a sufficient statistic for firms' extensive margin of trade for particular parametrizations of our model (which also turn out to coincide with the setup considered in Halpern, Koren, and Szeidl (2009) and Gopinath and Neiman (2012)), in general the theory implies that the identity of sourcing countries should be controlled for. In our second approach, we therefore use the entire variation in the firm-level data and allow for sourcing-strategy-specific fixed effects in firms' import demand system. Finally, we take an even more stringent approach and consider different

[4]Formally, the firm's maximization problem is non-convex and requires a search over all possible combinations of sourcing countries. Even with 10 products and 5 countries, the firm has a set of 50 varieties to choose from, which yields a total of 2^{50} sourcing strategies to evaluate.

[5]Note the importance of conditioning on the sourcing strategy. Gopinath and Neiman (2012) stress that their model features a non-homothetic import demand. This non-homotheticity, however, is entirely driven by, what they call, the sub-extensive margin of trade, where more productive firms source more products from abroad. Holding this margin fixed, their model is also homothetic.

[6]More precisely, we require that, conditional on importing a particular variety, all firms face the same price.

subsets of firms that share exactly the same sourcing strategy. All three approaches yield the same result: import demand is not homothetic. Moreover, the specific direction in which the theory is rejected is economically meaningful. We find that larger firms spend a higher proportion of their budget on their most important variety, after controlling for the sourcing strategy. Any model trying to be consistent with this salient feature of the firm-level data therefore has to introduce a mechanism that induces bigger firms to bias their import spending towards their most important trading partner, holding the extensive margin of trade fixed.

We discuss three particular mechanisms that can make the existing models of importing qualitatively consistent with the firm-level data. We first consider an extension where we do not restrict productivity to be factor-neutral. Using the information contained in firms' input prices, i.e. their unit values, we show that, after controlling for firms' extensive margin of trade, larger firms tend to spend a higher proportion of their budget on more expensive varieties. To the extent that expensive varieties feature high quality flows, this implies that variety quality and firm productivity are complements (Kugler and Verhoogen (2011, 2009)).[7] We then consider a simple process of search, where importers enter foreign markets and then have to search for a particular supplier. As more productive firms have higher incentives to search, they will - on average - receive better matches. Depending on the underlying distribution of match quality, this process will make the data look non-homothetic even though the actual production technology does not feature any bias. Finally, we discuss the importance of intra-firm trade. If FDI allows importing firms to save on unit prices (e.g. through lower variable trade costs or through improved product quality) and more productive firms are more likely to engage in FDI, the resulting demand system will also appear to be non-homothetic.

Related Literature. The vast majority of the recent contributions in trade have been concerned with the behavior of exporters and is summarized in Melitz and Redding (2012). A first generation of empirical papers established a number of stylized facts, namely that few firms engage in exporting, that exporters are larger and more productive than non-exporters, and that exporters usually sell most of their output domestically (Bernard and Jensen, 1999; Bernard, Jensen, Redding, and Schott, 2007, 2012). These findings were accompanied, on the theory side, by the development of frameworks that allow for firm-level heterogeneity and assume fixed costs to exporting (Melitz, 2003; Bernard, Eaton, Jensen, and Kortum, 2003; Chaney, 2008). More recently, Eaton, Kortum, and Kramarz (2004, 2011); Arkolakis and Muendler (2011) and Arkolakis (2010) have uncovered a number of novel facts regarding the entry behavior of firms into different export markets, and have subsequently shown how these facts can be theoretically rationalized by augmenting Melitz (2003) with various dimensions of firm-level heterogeneity. Arkolakis and Muendler (2011) and Bernard, Redding, and Schott (2011) explicitly introduce multi-product firms in the analysis and study market entry on the firm-product-country level.

[7] Additionally, we establish an asymmetry between the intensive and extensive margins of importing. On the extensive margin, we find that larger firms tend to enter more markets, with their marginal market being of lower quality, as measured by prices. This is consistent with the results on the export literature - see Eaton, Kortum, and Kramarz (2011). On the intensive margin, however, this relationship is reversed: holding their sourcing strategy fixed, larger firms spend more in expensive markets.

Compared to these developments on the export side, our knowledge about firms' import decisions is rather limited. Bernard, Jensen, Redding, and Schott (2007) show that the basic firm level facts about exporters are replicated for importers in the US, namely that importing is a rare activity and that importers are bigger and more productive than non-importers. More recently, various papers have taken a more structural approach to study the interaction between firms' importing decisions and productivity. Halpern, Koren, and Szeidl (2009) establish some new firm-level facts on entry into import markets and estimate a structural model of import behavior to identify and quantify the mechanisms driving the productivity gains associated with imported inputs. Gopinath and Neiman (2012) study the trade collapse during the Argentine crisis and stress the importance of firm-level heterogeneity, as the within-firm churning of imported products plays an important role in accounting for aggregate trade flows. Studying the trade liberalization in India, Goldberg, Khandelwal, Pavcnik, and Topalova (2010) also stress the productivity effects of imported varieties.[8] While they rely on reduced-form methods, they explicitly derive their estimation equations from a structural model of importing. This approach is also taken in Amiti, Itskhoki, and Konings (2012), who argue that the fact that most exporting firms are also importers can explain the exchange-rate disconnect. What all these papers have in common (despite their very different focus) is that they all share the same general framework, which we refer to as the benchmark model of importing. By testing the main robust prediction of this framework, we therefore test a property of the import demand system which all these approaches share.

Finally, our paper is also closely related to the quality-productivity complementarity hypothesis as stressed in Kugler and Verhoogen (2009, 2011). Using data from Colombia, they focus on domestic supplier relationships and show that more productive firms pay higher prices for their intermediate inputs. They offer two possible theoretical mechanisms to account for this finding: a complementarity between firm productivity and input quality, and the presence of fixed costs to acquire high quality inputs. We also consider the complementarity hypothesis as an explanation for our findings. However, there are two differences between our and their approach. First of all, we focus solely on imported inputs. Secondly, we show that bigger firms spend relatively more on expensive imports holding the extensive margin of importing constant. Thus, we are able to more clearly distinguish between technological complementarities and non-homotheticities induced by fixed costs. Relatedly, Manova and Zhang (2012) also show that exporters (i.e. bigger and more productive firms) source more expensive inputs from abroad.

The structure of the paper is as follows. We start in section 2 by describing a benchmark framework of import behavior, derive its empirical predictions and explicitly show how the existing contributions in the field are particular parametrizations. Section 3 contains the empirical analysis and proceeds in two steps. We first present two novel facts about the variety margin of importing which, through the lens of the theory, imply that different varieties within narrowly defined products are imperfectly substitutable and that price-adjusted qualities across countries are not equalized. We then test the implications of the theory for the intensive margin of importing and show that the

[8] At a reduced form level, an early paper about the firm-level productivity gains of importing is Amiti and Konings (2007).

predicted exclusion restriction is violated. In Section 4 we discuss three extensions of the model to reconcile the theoretical framework with the evidence. Section 5 concludes.

2 A Theoretical Framework of Import Behavior

In this section we present a general theoretical framework to account for firms' import demand. We model firms' import decisions as the solution to a static cost minimization problem. Production is subject to constant returns and firms are assumed to be price takers. We allow for unrestricted heterogeneity in firm productivity and fixed costs of importing, but neither prices nor input qualities are allowed to be firm-specific. As most existing contributions on firms' import behavior are special cases of this framework, we refer to it as the benchmark model of importing. The goal of this section is to derive strong testable predictions on the intensive margin and to show, along the way, that general predictions cannot be obtained for the extensive margin of import behavior.

2.1 The Environment

We assume that each firm has access to a production function

$$y = \varphi q(z), \tag{1}$$

where φ denotes the firm's factor neutral productivity (TFP), z is a vector of inputs and q is a constant returns to scale production function. For simplicity we suppress the explicit choice of other inputs like labor or capital. Inputs can either be sourced domestically or they can be imported. In particular, to keep the model close to the empirical analysis, we make the usual distinction between *products* and *varieties*. There is a set of products $\mathcal{K}$ (with n elements) and a set of countries $\mathcal{C}$ (with m elements), from which the foreign inputs 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 takes the following nested form:

$$q(z) = f(x_1, .., x_n) \text{ with } x_k = g_k(\eta_{1k} z_{1k}, ..., \eta_{mk} z_{mk}). \tag{2}$$

Here f and $\{g_k\}_k$ are constant returns production functions, η_{ck} parametrizes the quality of product k supplied by country c and z_{ck} denotes the amount of product k sourced from country c.[9] The main assumption encapsulated in the nesting structure of equation (2) concerns the degree of substitutability between different varieties. Crucially, the marginal rate of substitution between two varieties of a product k is given by

$$\frac{\partial q(z)/\partial z_{ck}}{\partial q(z)/\partial z_{c'k}} = \frac{\partial g_k(z)/\partial z_{ck}}{\partial g_{k'}(z)/\partial z_{c'k}} \frac{\eta_{ck}}{\eta_{z_{c'k}}}, \tag{3}$$

[9] With a slight abuse of notation, we define the production functions f and g_k over the entire sets $\mathcal{K}$ and $\mathcal{C}$ as we can always set $\eta_{ck} = 0$ if a particular product/variety is deemed technologically useless.

and hence does neither depend on the production function f nor on any allocations $(x_{k'})$ or technologies $(g_{k'})$ in different product classes $k' \neq k$.

Importing intermediate inputs from abroad is subject to both variable transport and fixed costs. In particular, we assume that firms take the set of prices $[p_{ck}]_{ck}$ as given and these prices contain all variable transport costs which accrue whenever a foreign variety is acquired, that is, $p_{ck} = (1+\tau_{ck})p^*_{ck}$, where p^*_{ck} is the price of product k in country c and τ_{ck} are the iceberg transport costs of product k from country c to France. The fixed costs of importing product k from country c are given by κ_{ck}. Additionally, there is a fixed cost of being an importer, which we denote by κ^I.

We will see below that the robust predictions of the theory concern the intensive margin of firms' import demand and follow from the solution to a cost-minimization problem. Hence we do not have to impose any structure on the demand side.

The crucial ingredients of this environment are as follows. First, we allow productivity φ and fixed costs $[\kappa_{ck}, \kappa^I]$ to vary across firms.[10] In contrast, variety qualities $[\eta_{ck}]$, prices $[p_{ck}]$ and variable transport costs $[\tau_{ck}]$ are assumed to be common to all firms but allowed to vary at the product-country level. We will see below that it is these two assumptions that give empirical content to the theory and imply share predictions for the intensive margin of trade.

2.2 Optimal Import Demand

In this section, we characterize firms' import demand system. To do so, it is useful to define a *sourcing strategy,* which describes firms' decisions as to which products to import and where to import them from - that is, firms' extensive margin of importing. More precisely, a sourcing strategy Σ is a subset of products $K \subset \mathcal{K}$ and subsets of countries $\{\Sigma_k\}_{k\in K}$ with $\Sigma_k \subset \mathcal{C}$, such that the firm imports positive amounts from these products and varieties.

It is convenient to split the firm's problem into the cost-minimization problem given a sourcing strategy Σ and the choice of the optimal firm size y and sourcing strategy given the cost function. Formally,

$$\pi \equiv \max_{\Sigma,y} \left\{ py - \Gamma(\Sigma, y, \varphi) - \sum_{(c,k)} \kappa_{ck} - 1(\Sigma)\kappa^I \right\}, \tag{4}$$

where

$$\Gamma(\Sigma, y, \varphi) \equiv \min_{z} \left\{ \sum_{(c,k)\in\Sigma} p_{ck} z_{ck} \text{ s.t. } \varphi q(z) \geq y \right\}, \tag{5}$$

is the firm's cost function. Here p denotes the demand function the firm faces, $1(\Sigma)$ is an indicator of import status and $\Sigma = [K, \{\Sigma_k\}_{k\in K}]$ is the firm's sourcing strategy.

[10] The fact that firms differ in their productivity level is a crucial component of the theory. In contrast, whether we assume that fixed costs are firm-specific or not turns out to be irrelevant for the focus of this paper, because we are mostly interested in the intensive margin of trade.

Firms' intensive margin of import demand is fully determined by the solution to (5). Let

$$\xi_{ck} \equiv \frac{\eta_{ck}}{p_{ck}} = \frac{\eta_{ck}}{(1+\tau_{ck})\, p^*_{ck}} \tag{6}$$

denote the quality flow per dollar spent on product k imported from country c. For brevity, we will refer to $\xi's$ as qualities.[11]

Letting $z_{ck}(\Sigma, y, \varphi)$ be the optimal amount sourced from country c of product k, it can be shown that[12]

$$s_{ck}(\Sigma, y, \varphi) \equiv \frac{p_{ck} z_{ck}(\Sigma, y, \varphi)}{\sum_{j \in \Sigma_k} p_{jk} z_{jk}(\Sigma, y, \varphi)} = h_k\left(\left[\xi_{ck}\right]_{c \in \Sigma_k}\right), \tag{7}$$

where h_k is a (product-specific) function that only depends on the qualities of the varieties sourced within the product, but not on the firm-characteristics (y, φ). Hence, (7) implies that conditional on the sourcing strategy Σ, the within-product allocation of expenditure is equalized across firms. Formally, letting $G^i_k(s)$ be firm i's distribution of expenditure on the different varieties sourced of product k, (7) implies that

$$G^i_k(s) = G_k\left(s|\Sigma^i_k\right), \tag{8}$$

i.e. the (endogenous) sourcing strategy (for product k) Σ^i_k is a sufficient statistic for firms' allocation of spending across sourcing countries so that G_k on the right-hand side of (8) ceases to have an i superscript.[13] In the logic of our theory, there is essentially a single source of heterogeneity - productivity. In that case, (8) can be written as $G_k\left(s|\Sigma^i_k, \varphi^i\right) = G_k\left(s|\Sigma^i_k\right)$, i.e. productivity has no effect on expenditure shares, other than through its effect on the set of varieties sourced and two firms who source product k from the same set of countries should have the exact same expenditure shares.

Testing this exclusion restriction will be at heart of the empirical analysis of this paper. That the intensive margin of import demand is homothetic follows from the assumptions that (i) TFP is factor neutral, and (ii) 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 (2). Together with the assumption that prices, qualities and variable 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.*[14]

[11] The inverse of ξ is sometimes referred to as pure prices - see for example Hallak and Schott (2011).

[12] See section 6.3 in the appendix for the formal derivation.

[13] While (8) is a necessary condition for (7), it is not sufficient. The reason is that (7) 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 (7) implies that all firms should spend the same share of their import budget on country A, (8) only implies that all firms should have the same allocation of spending across the two countries, irrespective of their identity. We not only think of (8) as the economically more important theoretical restriction to be tested, but we also revisit this issue in Section 3.4, where we show how an analysis based on (7) might lead to wrong conclusions.

[14] In fact, Halpern, Koren, and Szeidl (2009, 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

To actually implement (8), we are going to focus on firms' order statistics of their expenditure shares, i.e. their ranked expenditure shares. With a slight abuse of notation, let s^i_{jk} denote the j-th order statistics of firms $i's$ expenditure shares on product k.[15] Then, (8) implies the set of restrictions[16]

$$s^i_{jk} = s_{jk}\left(\Sigma^i_k, \varphi^i\right) = s_{jk}\left(\Sigma^i_k\right) \text{ for all } j = 1, ..., \left|\Sigma^i_k\right|. \tag{9}$$

To see the content of (9), it is useful to consider the case in which the production structure takes the CES form. In this case, we have that

$$\begin{aligned} f\left(x_1, ..., x_k\right) &= \left(\sum_{k\in\mathcal{K}} B_k x_k^{\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}} \\ x_k &= \left(\left(z^D_k\right)^{\frac{\zeta-1}{\zeta}} + \left(m^F_k\right)^{\frac{\zeta-1}{\zeta}}\right)^{\frac{\zeta}{\zeta-1}} \\ m^F_k &= \left(\sum_{c\in\mathcal{C}_k} \left(\eta_{ck} z_{ck}\right)^{\frac{\rho-1}{\rho}}\right)^{\frac{\rho}{\rho-1}}. \end{aligned} \tag{10}$$

Hence, each product-level composite x_k is a CES-aggregate of the domestic variety z^D_k and foreign input bundle m^F_k, where the bundle m^F_k itself aggregates foreign imports $[z_{ck}]$ in a CES way. Under these assumptions, firms' expenditure shares across foreign varieties in a given product are simply given by the CES analog to the general exclusion restriction (7), that is

$$s_{ck}\left(\Sigma, y, \varphi\right) = \frac{\xi^{\rho-1}_{ck}}{\sum_{j\in\Sigma_k}\xi^{\rho-1}_{jk}}. \tag{11}$$

Hence, the distribution in (8) fully inherits the properties of the distribution of $\xi^{\rho-1}$ and the j-th order statistic of expenditure shares s^i_{jk} is simply proportional to $\xi^{\rho-1}_{j,\Sigma_k}$, which is the j-th order statistic of $[\xi_{ck}]_{c\in\Sigma_k}$.

Given optimal import demand at the intensive margin, consider now the firm's choice of its sourcing strategy. As the production function features constant returns to scale, the cost function has a very simple form. In particular, it is given by

$$\Gamma\left(\Sigma, y, \varphi\right) = \gamma\left(\Sigma\right)\frac{1}{\varphi}y = \gamma\left(\left\{\left[\xi_{ck}\right]_{c\in\Sigma_k}\right\}_{k\in K}\right)\frac{1}{\varphi}y, \tag{12}$$

where $\gamma\left(\Sigma\right)$ is the unit cost function, which only depends on the sourcing strategy Σ. The second equality stresses that it is only the distribution of qualities $[\xi_{ck}]$ that determines firms' unit costs.[17]

margin of trade fixed, the cross-sectional variation in trade shares identifies firms' sourcing strategies.

[15] Using the usual notation, we should be writing $s^{(j),i}_k$, where (j) denotes the j-th order statistics. We use s^i_{jk} as a short-hand for this expression.

[16] As expenditure shares have to add up to unity, (9) is a set of $\left|\Sigma^i_k\right| - 1$ restrictions, which we are going to test.

[17] For a derivation of this result see Section 6.3 in the Appendix.

Furthermore, the unit cost function is weakly decreasing in Σ in the sense that $\gamma(\Sigma') < \gamma(\Sigma)$ whenever $\Sigma \subset \Sigma'$.[18] Given the cost function $\Gamma(\Sigma, y, \varphi)$ in (12), firms' import behavior then simply concerns the choice of the efficient firm size and the optimal sourcing strategy, i.e. the extensive margin of trade. These are defined by

$$(\Sigma^*, y^*) = \arg\max_{(\Sigma,y)} \left\{ \left[p - \gamma(\Sigma) \frac{1}{\varphi} \right] y - FC(\Sigma) \right\}, \tag{13}$$

where $FC(\Sigma) = \sum_{(c,k)\in\Sigma} \kappa_{ck} + 1(\Sigma)\kappa$ denotes the fixed costs of importing and p denotes the demand function.

The usual intuition of Melitz-type models of the exporting literature suggests that the extensive margin of import demand should satisfy a sorting condition with respect to firm productivity φ, i.e. not only import status but also the number of products and the number of varieties sourced should be positively correlated with firm productivity so that international sourcing should be hierarchical. This intuition, however, is incorrect. The reason is the interdependence between the different choices on the extensive margin. International sourcing on the input side is a vehicle to reduce the variable cost of production. Hence, a particular variety is imported whenever the reduction in the average production costs outweighs the incurred fixed costs. As long as there is some complementarity across imported varieties, i.e. as long as the production function features some form of "love for variety", these cost reductions depend on the *entire* sourcing strategy Σ. Thus, it might be that unproductive firms source multiple varieties with low fixed costs and low quality flows and high productivity firms concentrate on few fixed cost expensive varieties, which yield high quality flows. This interdependence renders the characterization of the extensive margin of importing much harder than for the case of exports. For exporting firms, the (outward) sourcing strategy can essentially be solved "market by market" - at least as long as production subject to constant returns to scale - see for example Eaton, Kortum, and Kramarz (2011). For imports, however, interdependencies in production are likely to be crucial. The intuition from the exporting literature that firms' extensive margin of trade features a hierarchy has therefore no direct counterpart for firms' importing decisions.[19]

Recent contributions are able to make progress on firms' extensive margin of trade - see Halpern, Koren, and Szeidl (2009) and Gopinath and Neiman (2012).[20] To do so, they make two crucial assumptions. First, they assume that foreign varieties *within* a product class are perfect substitutes, i.e. $\rho = \infty$.[21] Secondly, they assume that all imported products have the same price-adjusted

[18] This follows directly from the definition of Γ in (5) - given $\Sigma' \supset \Sigma$, $\Gamma(\Sigma, y, \varphi)$ was achievable so that $\Gamma(\Sigma', y, \varphi) \leq \Gamma(\Sigma, y, \varphi)$.

[19] 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 productive 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 productive firms source *more* varieties or products.

[20] A special case of their model is also analyzed in Ramanarayanan (2012). Goldberg, Khandelwal, Pavcnik, and Topalova (2010) use exactly our more general framework to study the effect a trade reform on firms' incentives to introduce new products. In their analysis they do not solve for firms' extensive margin of importing but only for the unit-cost function.

[21] To be precise, Gopinath and Neiman (2012) restrict their analysis to the case of a single good ($\mathcal{K} = 1$). Given the

quality, i.e. $\xi_{ck} = \xi$.[22] Under the additional restriction that fixed costs are constant across products ($\kappa_{ck} = \kappa$), it is possible to characterize firms' sourcing strategy almost analytically.[23] The reason is that, precisely under these assumptions, there is *no* interdependence of sourcing countries as there is only one "active" variety per product and the sourcing strategy reduces to a simple uni-dimensional choice: What is the optimal number of products to import? Thus, for these special cases, we recover the sorting result from the exporting literature that more productive firms access more markets because of the complementarity between φ and the unit costs $\gamma(\Sigma)$ (see (13)). Moreover, under this parametrization of the model, the theory implies that firms follow a perfect hierarchy in the sense that less productive firms source a strict subset of commodities of their more productive peers and that expenditure shares across imported inputs are equalized (see (11)).

In this paper we depart form these particular assumptions. We are able to do so because we will not focus on the extensive margin of trade, which relies crucially on these assumptions, but rather on an intensive margin property that holds in the general framework. Additionally, we will show that these assumptions are at odds with two novel firm-level facts that we report below. On the extensive margin we show that many firms do in fact source multiple varieties per product. This requires some complementarity between the different varieties, i.e. $\rho < \infty$. On the intensive margin we show that expenditure shares within firm-product cells across sourcing countries, far from being equalized, are highly concentrated on few suppliers. This is inconsistent with the assumption of prices, ξ_{ck}, being equalized.

3 Empirics

With this theoretical framework in mind, we now turn to the empirical evidence on the import behavior for the population of French manufacturing firms. We focus on the variety margin of international trade, where a variety is defined as an 8-digit product coming from a particular country. The main goal of this section is to provide a test for the robust intensive-margin prediction of the theory outlined above, namely, that the pattern of expenditure across varieties should be driven only by attributes of sourcing countries (i.e. qualities and prices) and not differ across firms once the extensive margin of trade is controlled for. Using different techniques we show that this assumption is soundly rejected in the data, and suggest, in the next section, three alternative mechanisms that can account for this result. Before turning to the test, we report in Section 3.2 two novel facts on firms' import behavior, which - through the lens of the model above - imply that both imperfect substitutability between varieties within products and price-adjusted quality differences across trading partners are important.

symmetric CES structure encapsulated in (10), this is equivalent to the assumption of there being multiple products and varieties within products being perfect substitutes.

[22] In fact, Halpern, Koren, and Szeidl (2009) also assume that the outer production function f takes the Cobb-Douglas form ($\varepsilon = 1$), where the product-wide expenditure shares B_k are allowed to vary.

[23] Relatedly, Lapham and Kasahara (2013) estimate a structural model of exporting and importing in a discrete choice framework. They can solve for the extensive margin of importing by assuming only a single foreign country, so that the interdependencies across international suppliers disappear.

	Quantiles				
	25%	50%	75%	90%	99%
No of annual trade interactions per country	10	53	328	3,763	69,371
No of annual trade interactions per product	8	25	74	179	752
No of annual trade interactions per firm	2	8	25	59	238
Value of firm-product-country interactions (1000EUR)	44.15	268.53	1,258.83	5,319.10	70,993.13
Number of firm-product-country observations			705,316		

Notes: See Section 6.1 in the Appendix for a description of the data.

Table 1: The Concentration of French Imports

3.1 Data and Descriptive Statistics

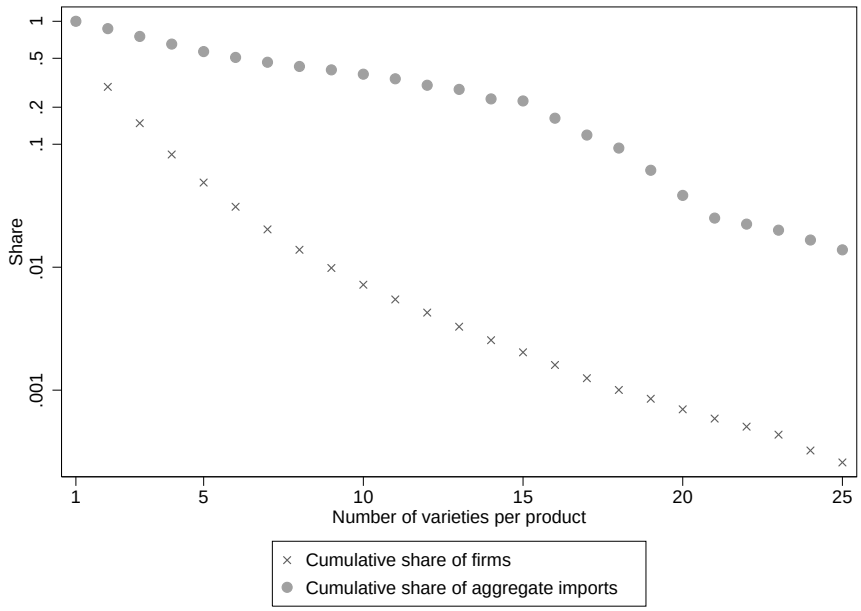
In this section we provide a general overview of the dataset. A detailed description of how the data is constructed is contained in the Appendix. Because we are interested in the demand for inputs, we restrict the analysis to manufacturing firms. We observe import flows for every manufacturing firm in France from the official custom files. Manufacturing firms account for 31% of the population of French importing firms and 56% of total import value in 2001. Overall, French firms trade with a total 226 countries. The flows are classified at the 8-digit (NC8) level of aggregation, which means that the product space consists of roughly 9,500 products. Using unique firm identifiers we can match this dataset to fiscal files, which contain detailed information on firm characteristics. The final sample consists of an unbalanced panel of roughly 260,000 firms which are active between 2001 and 2006.[24]

As the existing literature did not focus on the variety margin of firm-level trade data, we summarize the distribution of variety flows, i.e. product-country cells, across firms in Table 1 .[25] In total we observe roughly 700,000 variety-firm pairs. Given that there are about 30,000 importers in our data, the average importer imports about 23 varieties of potentially different products. Table 1 however shows that this average is not too informative as international activity is highly concentrated, both geographically (row one) and in the product space (row two). The median country is only active in 50 firm-product cells, whereas the top two exporting countries to France, namely Germany and Italy, report 70,000 interactions in distinct firm-product cells. Similarly, for half of the potential products, i.e. roughly 5,000 products, only 25 country-firm interactions are observed, while the most popular products are shipped into France in more than 750 distinct country-firm combinations.[26] Finally, the two remaining rows confirm the findings of Gopinath and Neiman (2012) 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 (that is, 300 firms) import 240 varieties. Similarly, while the most active firm-variety pairs are worth more than 70m EUR, a quar-

[24]This dataset is not new and has been used in the literature before (e.g. Mayer, Melitz, and Ottaviano (2010); Eaton, Kortum, and Kramarz (2004, 2011)). However, these contributions focused almost exclusively on the export side.

[25]Table 1 is of a similar flavor as the discussion of sparsity by Armenter and Koren (2008) but it is slightly different. Whereas they analyze the data on the *flow* level, we aggregate the data within a *firm-variety cell*, because it is these quantities which our model can speak to.

[26]Note in particular the still very large difference between the 99% and 90%-quantile.



Notes: The figure shows the the share of importers, who source at least C_i varieties per product and the share of aggregate imports these firms account for. C_i is the (expenditure share weighted) average of the number of varieties per product, i.e. $C_i = \sum_k s_{k,i} C_{k.i}$, where $s_{k,i}$ is the expenditure share of firm i on product k and $C_{k,i}$ is the number of countries firm i sources product k from. We use 6 years of data from 2001-2006 and report the yearly average.

Figure 1: How Many Varieties Do Firms Source?

ter of French importers import less than 45,000EUR worth of the varieties within a year. Thus, at whatever dimension we look, the world of imported inputs is a small world. Few firms are actively participating, and when they do, they tend to import from a small set of countries and only a small subset of potential commodities.

3.2 The Variety Margin of Importing: Two Facts

In this section we report two, to the best of our knowledge, novel facts about firms' demand for varieties - both on the extensive and the intensive margin.

Fact 1. There is substantial heterogeneity in firms' variety sourcing behavior at the extensive margin. 80 percent of the firms source at most 3 varieties per product, while 1 percent source more than 8. Also, larger firms tend to import more varieties per product. In Figure 1 we depict the distribution of the average number of varieties per product at the firm level. Specifically, we show the share of firms who import at least C varieties per product for different values of C. We also depict the share of aggregate imports, these firms account for. The Figure shows substantial heterogeneity in firms' sourcing behavior: while the vast majority of firms sources a limited number of varieties - 80 percent of French firms source less than 3 varieties per product - there is a group of firms that sources many - about 1 percent of the firms source more than 8 varieties per product. Moreover, it is these (few) firms that determine aggregate trade flows as they

	Number of Varieties Per Product						
	1	2	3	4	5	6	>10
Employment	23.253	73.436	121.194	170.853	231.713	341.71	1219.704
Sales	3774	14883	28386	45121	61703	101668	644222
Nb imported products (8 dig.)	3.91	18.32	27.523	32.909	38.57	48.706	76.804
Capital per worker	37.582	45.872	50.786	50.957	52.912	55.442	66.221
Share of exporters	0.58	0.81	0.883	0.923	0.949	0.97	0.996

Notes: The table shows different firm characteristics as a function of the average number of varieties sourced. See Section 6.1 in the Appendix for a description of the data.

Table 2: Cross-sectional Heterogeneity and Variety Sourcing

account for almost 50% of aggregate imports. This is also seen in Table 2, which displays the cross-sectional corollaries of the number of varieties sourced. It is clearly seen that the number of varieties sourced is positively correlated with firm size, as measured by sales or employment, capital-intensity, the number of products sourced and export status.[27]

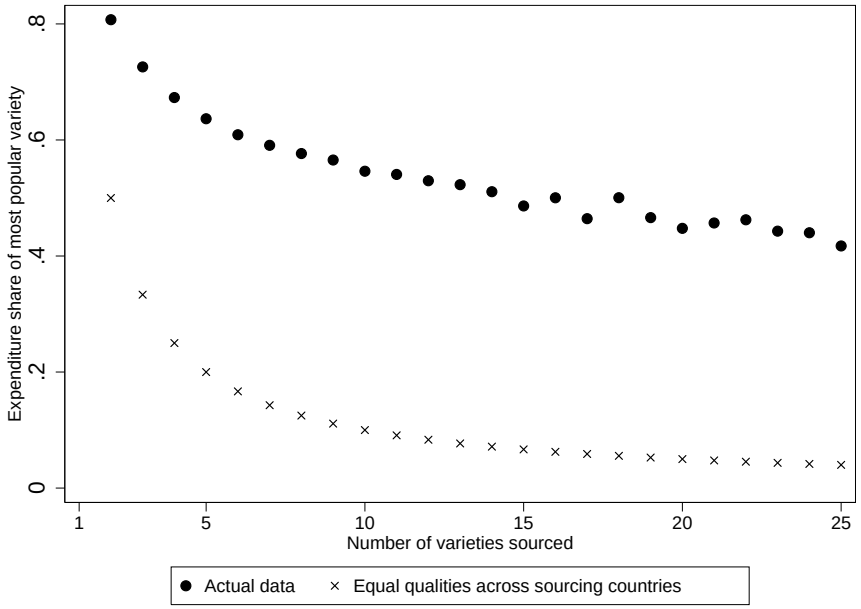
Fact 2. At the intensive margin, firms' sourcing behavior is highly concentrated. Firms importing 25 varieties per product spend 80 percent of the product's expenditure on 5 varieties. We now look at the allocation of expenditure across varieties within products. In Figure 2 we depict the average expenditure share on the main variety for firms sourcing V varieties per product for different values of V. Even firms sourcing 10 varieties spend more than 50 percent of their total expenditure on their main variety. 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 "natural" suppliers for most of their import budget, while many marginal sourcing countries seem to play a minor role.

Facts 1 and 2 are not only of interest in itself but, as argued above, crucially determine two key ingredients of theory.[28] Consider first the evidence on the extensive margin, i.e. Figure 1. Within the theoretical framework laid out above, complementarities across different varieties within a product are necessary to generate a pattern as seen in the data. For the special case of the CES model, Figure 1 implies that $\rho < \infty$. Similarly, consider the intensive margin of trade, depicted in Figure 2. According to the theory, firms' expenditure shares are fully determined by varieties' price-adjusted qualities ξ_{vk} (see for example (11)). Fact 2 therefore implies that, within a product k, prices do vary across supplying countries - i.e. $\xi_{vk} = \xi_k$ for all k does not hold.[29] While $\rho < \infty$

[27] That this cross-sectional variation is not merely due to technological differences is shown in Table 9 in the Appendix, where we show that there is still ample cross-sectional variation within industries.

[28] Both facts have a close counterpart on the exporting side. On the extensive margin (Fact 1), multi-market exporters also account for the largest share of aggregate exports (Arkolakis and Muendler, 2011; Bernard, Jensen, Redding, and Schott, 2012). On the intensive margin (Fact 2), multi-product firms also generate most of their sales from their top product (Arkolakis and Muendler, 2011; Bernard, Redding, and Schott, 2010).

[29] Note also that Figure 1 stresses the importance of using firm level data to study the variety margin of trade. Measuring the number varieties using aggregate data would suggest that the representative French manufacturing firm imports its intermediate inputs from more than 30 countries (because different firms do not necessarily agree on the identity of their sourcing countries). Figure 1, however, shows that this would be a poor representation of firms' import behavior as there are very few firms that source from more than 5 countries.



Notes: The figure shows the average expenditure share on the 5 most popular varieties of firm-product pairs with C varieties For the unweighted average all firms get an equal share, for the weighted results, we weigh firms by their import value. We also depict the counterfactual expenditure share if price-adjusted quality flows (and hence expenditure shares) were equalized across varieties.

Figure 2: Concentration of Firms' Import Spending Across Varieties

and non-constant ξ_{vk}'s seem to be natural cases to consider, we argued above that they preclude an analytic characterization of firms' sourcing strategy (i.e. of the extensive margin of importing) and are therefore violated in Halpern, Koren, and Szeidl (2009) and Gopinath and Neiman (2012). Our focus on the intensive margin allows us to bypass such restrictive assumptions.

3.3 Model Testing

The robust predictions of the basic theoretical framework concern the intensive margin of trade, namely that expenditure shares should not vary with firm characteristics once the sourcing strategy is controlled for.[30] Testing for this exclusion restriction amounts to a test of the joint hypothesis of productivity differences being factor neutral, and qualities and prices being common across firms.[31] As these assumptions are present in virtually all contributions modeling import demand, this implication tests an important pillar of the current theoretical structure in applied work. Furthermore, by testing this prediction, we are also able to assess whether characteristics of the sourcing country, such as transport costs, quality or prices are correlated with particular firm attributes. This will be informative on how to refine the theoretical framework to model import demand - see Section 4 below.

As explained above, we are going to test for non-homotheticities by focusing on the order statistics of firms' expenditure shares. In (9), we showed that the theory implies that

$$s_{jk}^{i} = s_{jk}\left(\Sigma_k^i, \varphi^i\right) = s_{jk}\left(\Sigma_k^i\right) \text{ for all } j = 1, ..., \left|\Sigma_k^i\right|, \tag{14}$$

where again Σ_k^i is the set of countries firm i sources product k from and s_{jk}^i denotes the j-th order statistic of firm i's expenditure shares (on product k). Considering a log-linear approximation of (14), we arrive at

$$ln\left(s_{jk}^{i}\right) = \alpha_{jk} + \phi\left(\Sigma_k^i\right) + \beta ln\left(\varphi^i\right), \tag{15}$$

where (14) implies that $\beta = 0$.[32] Note that (15) is exactly correct if the production function within products takes the CES form.[33]

The key to test for the exclusion restriction contained in (15) lies on how to control for $\phi\left(\Sigma_k^i\right)$.

[30] While the general theory presented in Section 2 features no strong predictions on the extensive margin, we looked at these partial correlations in a regression framework. As in Bernard, Jensen, Redding, and Schott (2007), Halpern, Koren, and Szeidl (2009) and Kugler and Verhoogen (2009), larger firms are more likely to be importers and they import more products. Additionally, and consistent with Table 2, larger firms also import more varieties per product. For completeness these results are contained in Table 10 in the Appendix.

[31] In general, the assumption of constant returns to scale is also required for the exclusion restriction to hold. In the CES case, it turns out that constant returns are not required.

[32] Recall that it is the nesting property of the production function (3) that allows us to consider expenditure shares product-by-product.

[33] From (11) it follows that

$$ln\left(s_{jk}^{i}\right) = (\rho - 1)\, ln\left(\xi_{j,C_k}\right) - ln\left(\sum_{j \in \Sigma_k^i} \xi_{jk}^{\rho-1}\right) \equiv (\rho - 1)\, ln\left(\xi_{j,\Sigma_k^i}\right) + \phi\left(\Sigma_k^i\right),$$

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

To make (15) operational we therefore consider three complementary approaches. First, we proxy the sourcing strategy Σ_k^i by its cardinality $\left|\Sigma_k^i\right|$, i.e. the number of varieties firm i sources from product k. While the number of sourcing countries is indeed a sufficient statistic for firms' extensive margin of trade for particular parametrizations of the model (Halpern, Koren, and Szeidl (2009) and Gopinath and Neiman (2012)), in general the theory implies that the identity of sourcing countries matters. Therefore, in our second approach we directly implement (15) by including a complete set of sourcing-strategy-specific fixed effects. Third, in our most stringent specification, we allow all coefficients to vary with the sourcing strategy.

Approach 1: Proxying firms' sourcing strategy by its cardinality. We now implement (15) by controlling for $\phi\left(\Sigma_k^i\right)$ by the number of countries product k is sourced from, $\left|\Sigma_k^i\right|$. We view this step as an approximation since, in general, the identity of the countries in the sourcing strategy Σ_k^i may matter[34] - the next subsection addresses this issue. To maximize the number of observations, we focus on the maximum and the minimum expenditure shares[35], which are defined for all firms sourcing at least two varieties.[36] Hence, we pool the data and estimate the regression:

$$ln\left(s_{max,k}^i\right) \quad = \quad \alpha_k - \gamma ln\left(\left|\Sigma_k^i\right|\right) + \beta ln\left(S^i\right) + \mu X^i + u_k^i, \tag{16}$$

where $s_{max,k}^i = \max_{c\in\Sigma_k^i} s_{ck}^i$ is firm i's expenditure share on its most important variety of product k, α_k is a product-specific intercept, $\left|\Sigma_k^i\right|$ denotes the number of varieties firm i sources of product k, X^i are firm-specific characteristics which control for differences in technology and S^i is firm sales, which we take as a proxy for productivity (the sole source of variation in the theory).[37]The theory of Section 2 implies $\beta = \mathbf{0}$. While the form of the regression in (16) is exactly correct for the CES case with constant price adjusted qualities, the exclusion restriction itself does not rely on these assumptions so that we interpret (16) as a linear approximation to the non-parametric exclusion restriction of the general case - see (7). Panel A of Table 3 reports the results.

The first two columns of Panel A in Table 3 show that more productive firms, as measured by sales, and firms sourcing more varieties tend to spend a lower fraction of their product expenditure on their top variety. These results are not surprising. That firms sourcing more varieties per product spend less on their most popular one is almost mechanical. The negative coefficient for sales in column (1) reflects the positive correlation between firm productivity and the number of varieties

[34]We note, however, that under the assumptions that quality-adjusted prices are equalized across countries (i.e. $\xi_{ck} = \xi_k$), and fixed costs are not firm-specific, which are for example imposed in Halpern, Koren, and Szeidl (2009) and Gopinath and Neiman (2012), the number of countries is indeed a sufficient statistic for the sourcing strategy.

[35]In our second approach below, we consider the entire distribution of expenditure shares at the firm-product level.

[36]Clearly we need to exclude firm-product pairs that are sourced from a single country, as they feature expenditure shares that are by construction equal to unity.

[37]In the Appendix we also report the results when we consider estimated productivity as the dependent variable. Our data does not have information on physical output. Hence, productivity is a revenue-based measure. For this measure to be a meaningful proxy for physical productivity φ, we would have to impose particular assumptions on the market structure of output markets and specify if there are any constraints on firms' input choices. We believe that sales is therefore a more robust proxy for productivity as the source of variation. Nevertheless, estimated revenue is considered in numerous contributions (e.g. Bernard, Jensen, Redding, and Schott (2007)) and hence we report the results in Section 7.1 of the Appendix for completeness.

Panel A	Dep. Variable: Max. expenditure share $ln(s^i_{max,k})$				
	(1)	(2)	(3)	(4)	(5)
ln Sales	-0.003***		0.011***	0.009***	0.011***
	(0.000)		(0.000)	(0.000)	(0.001)
ln Nb. Varieties (prod.)		-0.279***	-0.289***	-0.289***	-0.294***
		(0.001)	(0.001)	(0.001)	(0.002)
Exporter				0.009***	0.007***
				(0.001)	(0.002)
Capital per worker				0.002***	0.003***
				(0.000)	(0.001)
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
N	655,648	655,648	655,648	594,903	594,903
R^2	0.06	0.22	0.23	0.23	0.61

Panel B	Dep. Variable: Min. expenditure share $ln(s^i_{min,k})$				
	(6)	(7)	(8)	(9)	(10)
ln Sales	-0.277***		-0.125***	-0.112***	-0.158***
	(0.002)		(0.001)	(0.002)	(0.004)
ln Nb. Varieties (prod.)		-3.337***	-3.230***	-3.219***	-2.860***
		(0.006)	(0.006)	(0.006)	(0.012)
Exporter				-0.056***	-0.035**
				(0.009)	(0.016)
Capital per worker				-0.018***	-0.030***
				(0.003)	(0.005)
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
N	655,648	655,648	655,648	594,903	594,903
R^2	0.13	0.44	0.45	0.45	0.77

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. We weigh observations such that each firm has an equal weight. In Panel A, the dependent variable is $ln\left(s^i_{max,k}\right)$, where $s^i_{max,k} = \max_{c \in C^i_k} s^i_{ck}$ is the expenditure share on the most popular variety of product k for firm i. In Panel B, the dependent variable is $ln\left(s^i_{min,k}\right)$, i.e. the expenditure share on the least popular variety. The number of varieties is the number of countries where product k is sourced from. Columns (5) and (10) include a full set of interacted fixed effects at the product-country-industry level. Columns (4), (5), (9) and (10) control for indicator variables if the firm is member of a foreign or corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment.

Table 3: Firm Characteristics and the Intensive Margin of Trade

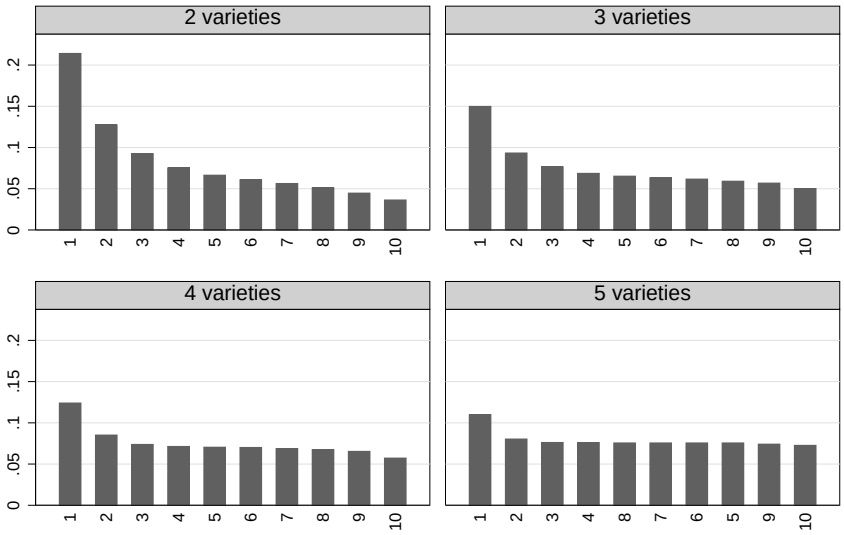
sourced. Our main interest is in column (3), where we find that *conditional on the number of varieties sourced*, firm sales has a positive and significant effect on the top variety expenditure share. This finding is at odds with the main prediction of the theory outlined in Section 2, according to which the effect of productivity should take place only via the sourcing strategy. Column (4) shows that these results are robust to including additional firm characteristics as exporting status, capital intensity and membership in a foreign or corporate group.[38] We see that the coefficients on sales and the number of varieties, as well as the R^2, hardly change after including the controls. To get a sense of the order of magnitude, the coefficient on sales reported in columns (3)-(4) implies that an increase in sales by one standard deviation increases the share of the top variety by 2%. Given the high concentration of expenditure shares, this corresponds also to an increase of roughly two percentage points. In column (5) we replace the product fixed effect α_k (see (16)) by $\alpha_{k,c(i),s}$, where k denotes the product, $c(i)$ the identity of firm i's most popular variety and s the industry of firm i. By doing so we control both for the fact that the number of varieties sourced is an imperfect control for the extensive margin and allow for industry-specific differences in varieties' quality (ξ_{ck}). While this increases the predictive power of the regression substantially, the coefficient on sales is unaffected. In sum, we find that more productive firms spend relatively more on their most popular sourcing country.

In Panel B of Table 3 we report the analog of regression (16), where we take firms' expenditure share on their least important variety, i.e. $ln\left(s^i_{min,k}\right)$, where $s^i_{min,k} = \min_{c\in\Sigma^i_k} s^i_{ck}$, as the dependent variable. It is clearly seen that we can again comfortably reject the prediction of the theory that β is equal to zero. Additionally, we now find that bigger firms spend less on their marginal variety. This of course is not entirely unexpected given the finding in Panel A, as firms' expenditure shares are by construction negatively correlated and many firms source only 2 varieties per product (see Figure 1). Finally, Table 13 in the online appendix shows that the results of Table 3 are robust to using different measures of productivity.

As explained above, (16) follows from (15) only under the stringent assumption that the number of sourcing countries appropriately controls for a firm's sourcing strategy. We now assess how reasonable this assumption is. Figure 3 below depicts the distribution of sourcing strategies for all firms sourcing the exact same number of varieties per product. Consider for example the first 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, calculate the share of firms in each of them, rank the different sourcing strategies by their popularity and average these distribution across products. The remaining panels of Figure 3 show the results of this exercise for case of 3,4 and 5 varieties per product. In all four cases, we find considerable heterogeneity in firms' extensive margin behavior. The histograms are far from degenerate, which would correspond to the case of perfect agreement. Within narrowly defined product classes, firms that agree in the number of varieties sourced exhibit considerable disagreement as to the identity of those varieties.[39] Thus, the assumption embedded

[38] In the interest of space, the estimated coefficients for membership in foreign group and membership in corporate group are not reported in Table 3 nor in the following Tables. They are available upon request.

[39] Note that this disagreement is *not* inconsistent with the theory. It is only in the very special, where quality



Notes: The graph plots the distribution of sourcing strategies, averaged across products, for sourcing strategies that involve exactly 2, 3, 4 or 5 varieties, for 2005. For each product with 10 firms or more, we select all firms sourcing the product from exactly v countries. We classify these firms according to their exact sourcing strategy. We rank the different sourcing strategies by popularity (where 1 denotes the most popular strategy) and average the resulting distributions across products. This procedure is done for v=2,3,4 and 5.

Figure 3: Disagreement on the Extensive Margin

in (16) that the number of sourcing countries is a sufficient statistic for a firm's sourcing strategy is problematic. Therefore, we now consider an alternative approach that controls for firms' sourcing strategy in a more rigorous way.

Approach 2: Explicitly controlling for firms' sourcing strategy. As discussed above and seen in (15), the theory requires us to control for the entire set of sourcing countries and not only for its cardinality. In other words, not only the number but also the identity of the countries matters. Hence we now implement equation (15) by introducing sourcing-strategy-specific fixed effects.

For comparability with the previous approach, we start by regressing the expenditure share of the top variety on sales and other firm characteristics, as well as on sourcing-strategy fixed effects. Columns (1) and (2) of Table 4 contain the results. Column (1) shows that the coefficient of sales is positive and significant. This suggests that productivity has a positive effect on the top variety expenditure share, after perfectly controlling for the sourcing strategy. Column (2) shows that this result is robust to including additional firm-level controls (the firm's export status, capital per worker, and indicator variables for membership in a corporate or foreign group).

Next, we focus on the entire distribution of firms' expenditure shares across sourcing countries. As expenditure shares add up to unity, we consider *relative* expenditure shares and run the following regressions:

$$ln(s^i_{jk}/s^i_{1k}) = \alpha_k + \alpha_\Sigma + \beta ln(S^i) + \mu X^i + u^i_{jk} \text{ for } j = 2, 3, ..V \tag{17}$$

where V is the number of varieties sourced by the firm-product pair, s^i_{jk} is the j-th order statistic of firm i's expenditure shares on product k, and α_Σ is a sourcing-strategy-specific fixed effect. As stressed above, we use the firm-specific ranking of expenditure shares and work with order statistics.[40] The regression in (17) tests the basic theory in its purest form. By comparing the distribution of expenditure shares of firms who source their inputs from exactly the same countries within 8-digit products, we focus precisely on the robust prediction of the theory: the intensive margin of trade should not depend on firm-characteristics, i.e. β should be zero in all specifications. Again, we want to stress that this implication does not rely on the CES functional form.[41] Columns (3)-(10) in Table 4 contain the results. Columns (3)-(4) show that the coefficient on sales is negative and statistically significant, both with and without additional firm controls. Note that the dependent variable, $ln(s^i_{2k}/s^i_{1k})$, is the log difference between the min and the max share. Thus, a negative coefficient means that larger firms bias their expenditure towards their most important variety. This is entirely consistent with the results in Table 3 above. Quantitatively, a 10% increase in sales increases the expenditure share on the top variety relative to the second most important variety by 1%. Columns (5)-(8) confirm these results for the cases where $ln(s^i_{3k}/s^i_{1k})$ and $ln(s^i_{4k}/s^i_{1k})$ are used as

differences are absent and fixed costs do not vary at the firm-level, that we would expect hierarchical sorting. See also the discussion on page 10.

[40] While in principle we could keep the identity of the varieties fixed within each regression, we choose to use a firm-specific ranking to deal with the fact that firms may disagree in their ranking of varieties by expenditure shares. We will come back to this point in Section 3.4 below.

[41] By choosing the country with the highest expenditure share as the benchmark, we are able to interpret the results as suggestive of a complementarity between firm productivity and input quality. See the discussion below.

	Dependent variable:									
	$ln(s_{ik}^{max})$		$\ln(s_2/s_1)$		$\ln(s_3/s_1)$		$\ln(s_4/s_1)$		$\ln(s_5/s_1)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln Sales	0.009***	0.009***	-0.097***	-0.092***	-0.145***	-0.140***	-0.152***	-0.133***	-0.128	-0.113
	(0.000)	(0.000)	(0.003)	(0.004)	(0.011)	(0.014)	(0.033)	(0.042)	(0.088)	(0.105)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Fixed Effects	Product-Sourcing Strategy									
N	655,648	594,903	655,648	594,903	262,691	239,225	130,951	119,862	73,534	67,659
R^2	0.73	0.74	0.58	0.59	0.83	0.83	0.93	0.93	0.97	0.97

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects, 8 digit product fixed effects and fixed effects for individual sourcing strategies. We weigh observations such that each firm has an equal weight. In columns 1 and 2, the dependent variable is $ln\left(s_{max,k}^{i}\right)$, where $s_{max,k}^{i} = \max_{c \in C_k^i} s_{ck}^{i}$ is the expenditure share on the most popular variety of product k for firm i. In the remaining columns, the dependent variables are $ln\left(s_j^i/s_1^i\right)$, which is the expenditure share on firm i's jth variety relative to its first variety, where the different varieties are ranked by their expenditure shares. All columns with controls control for firms' export status, capital per worker and for indicator variables if the firm is member of a foreign or corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment.

Table 4: Firm Characteristics and the Intensive Margin of Trade: Sourcing Strategy Fixed Effects

dependent variable. Columns (9)-(10) show that the results for the case of 5 varieties are statistically weaker as we lack power to estimate these effects precisely.

Approach 3: Sourcing-Strategy-Specific Coefficients In the previous approach we imposed common coefficients for sales across sourcing strategies. We now relax this restriction and allow all coefficients in (17) to vary by sourcing strategy. This amounts to estimating (17) on subsamples of firms that share exactly the same sourcing strategy. In this sense, this is the most stringent approach we employ to test the exclusion restriction implied by the theory of Section 2.

Table 5 contains the results. For expositional simplicity, we display the coefficients for sales that correspond to the most popular sourcing strategy by product.[42] Panels A, B, and C deal with firm-product pairs that source exactly 2,3 and 4 varieties, respectively. We confirm the results of the previous section: all coefficients are negative and statistically significant.[43] Furthermore, the coefficients are monotonically decreasing, which implies that more productive firms shift their entire distribution of expenditure towards their preferred sourcing countries. Formally, the distribution of expenditure across ranked sourcing countries of large firms first-order stochastically dominates the one of their smaller sized counterparts. Panel D shows that, as in Table 4, the results for the case of 5 varieties are statistically weaker as we lack power to estimate these effects precisely. While all coefficients are negative and monotone, many are not significant at conventional levels. We tackle this issue in Section 6.5 in the Appendix, where we make the additional assumption that technologies take the CES functional form. Putting this additional structure allows us to substantially increase sample size. The point estimates hardly change but are now statistically significant.

Discussion. The main takeaway of the results in Tables 3, 4 and 5 is the lack of empirical support for the central prediction of the theory outlined in Section 2. More precisely, firm-level productivity, as measured by sales, has a significant effect on expenditure shares after controlling for the sourcing strategy. This means that we reject the hypothesis that import demand is homothetic.

The results of this section also suggest a particular direction for the effect of firm-productivity on expenditure shares. In particular, all tables feature more productive firms spending *relatively more* on their top varieties. By controlling for firms' sourcing strategy, this pattern is a property of the firms' intensive margin of trade. As expenditure shares are increasing in price-adjusted input quality, these results suggest that large firms bias their expenditure towards varieties of high quality. In Section 4 below we propose three alternative mechanisms that can give rise to such bias.

[42] The algorithm we use is as follows. Consider the case of two countries. For each of the products, we select the two countries which, taken together, constitute the most popular two-variety sourcing strategy. We then keep all firms that source this product from exactly these two countries. 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. The fact that the sourcing strategy needs to be controlled for only at the firm-product level is an implication of the nesting property of the production function. We repeat this procedure for the cases of three, four and five sourcing countries.

[43] With the exception of column (2) in Panel C.

	$\ln(s_2/s_1)$		$\ln(s_3/s_1)$		$\ln(s_4/s_1)$		$\ln(s_5/s_1)$	
Panel A	Firms sourcing the exact same 2 varieties of product k							
ln Sales	-0.099***	-0.095***						
	(0.005)	(0.006)						
Firm Controls	No	Yes						
N	79,593	71,767						
R^2	0.24	0.25						
Panel B	Firms sourcing the same exact 3 varieties of product k							
ln Sales	-0.088***	-0.082***	-0.198***	-0.203***				
	(0.012)	(0.014)	(0.016)	(0.019)				
Firm Controls	No	Yes	No	Yes				
N	23,162	20,917	23,162	20,917				
R^2	0.50	0.51	0.54	0.54				
Panel C	Firms sourcing the exact same 4 varieties of product k							
ln Sales	-0.044*	-0.028	-0.098***	-0.081**	-0.208***	-0.192***		
	(0.025)	(0.028)	(0.031)	(0.038)	(0.040)	(0.049)		
Firm Controls	No	Yes	No	Yes	No	Yes		
N	10,764	9,707	10,764	9,707	10,764	9,707		
R^2	0.75	0.76	0.76	0.76	0.76	0.77		
Panel D	Firms sourcing the exact same 5 varieties of product k							
ln Sales	-0.046	-0.089	-0.066	-0.108	-0.123*	-0.133	-0.162*	-0.164
	(0.047)	(0.055)	(0.064)	(0.080)	(0.072)	(0.088)	(0.083)	(0.107)
Firm Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	6,486	5,862	6,486	5,862	6,486	5,862	6,486	5,862
R^2	0.87	0.88	0.87	0.88	0.88	0.88	0.87	0.87

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. We weigh observations such that each firm has an equal weight. The dependent variables are $ln\left(s_j^i/s_1^i\right)$, which is the expenditure share on firm i's jth variety relative to its first variety, where the different varieties are ranked by their expenditure shares. The table contains 4 sets of regressions, each corresponding to a different number of varieties in the sourcing strategy. In any regression, all firms share the exact same sourcing strategy for a given product (see main text for description of this procedure). All regressions that contain firm controls, control for firms' export status and (log) capital intensity and contain indicator variables if the firm is member of a foreign or a corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment.

Table 5: Firm Characteristics and Relative Expenditure Shares: Sourcing-Strategy-Specific Coefficients

	Mean	p25	p50	p75
σ_k	0.2720331	0.1775188	0.2748849	0.355045
δ_k	.4175134	.2592636	.4158728	.5632002
α_k	.6752327	.5	.6666667	.8333334
$\tilde{\sigma}_k$	0.1777728	0.0957896	0.1818412	0.2543983

Notes: The table displays summary statistics of the distribution of $\sigma_k, \alpha_k, \delta_k, \sigma_k$ across products. For each product, the corresponding dispersion measure is computed on a subsample of firms that agree on a 2-variety sourcing strategy.

Table 6: Disagreement on the Intensive Margin: Descriptive Statistics

3.4 Disagreement On The Intensive Margin

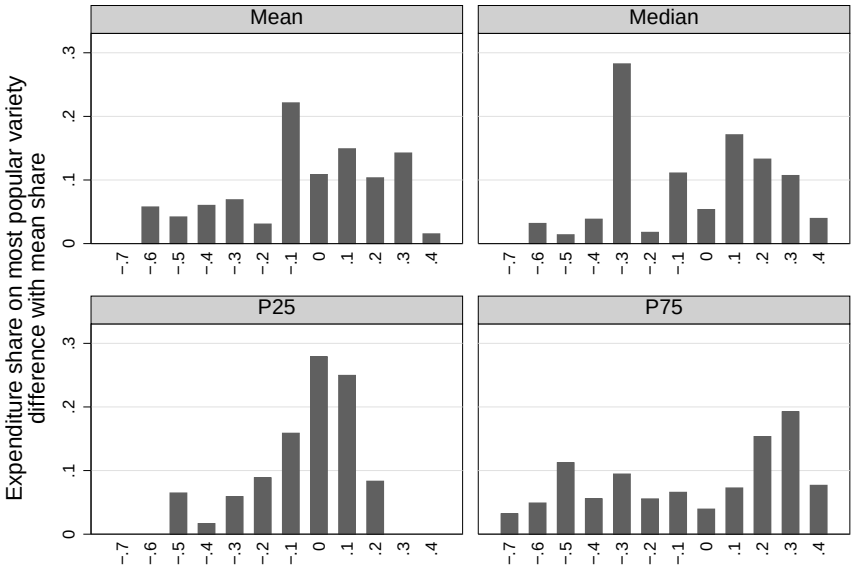
In the previous section, we focused on order statistics of expenditure shares to test for the homotheticity of the demand system. Using order statistics amounts to studying the firm-level distributions of expenditure shares, ignoring the identity of the country associated with each share. In terms of the theory of Section 2, we worked with equation (8), which is a weaker form of equation (7). Indeed, when prices, qualities and variable trade costs are assumed to be common across firms, the theory implies that expenditure shares *on a particular variety* should not vary at all among firms that share a sourcing strategy. In this section, we bring this strong prediction to the data and study the distribution of expenditure shares when country identities are taken into account. While we do not expect to find perfect equalization of shares due to the presence of measurement error or potential idiosyncratic factors, we nevertheless assess the degree of variation in the intensive margin of trade.[44]

To get a quantitative measure of the degree of agreement in the intensive margin of trade, we adopt the following procedure. We use the subsamples constructed in Approach 3 above and focus on the case of 2 varieties. For each product k, and associated sourcing strategy, we select the country that features the largest share of firms sourcing it as their main variety (call it country M_k) and compute the distribution of expenditure shares on this country across firms ($s_{M_k,i}$). We then measure, for each product, the dispersion in shares across firms with the following four statistics: (i) the standard deviation, σ_k, (ii) the ratio of 75th to 25th percentile, δ_k, (iii) the share of firms that feature country M_k as their main variety[45], α_k, (iv) the residual standard deviation of expenditure shares after controlling for sector fixed effects[46], $\tilde{\sigma_k}$. Table 6 summarizes these product-specific dispersion measures across products. We find considerable variation in expenditure shares, with an average standard deviation of 0.27 and an average 75-25 percentile ratio of 0.41. We also find that, on average across products, only 67% of firms feature country M_k as their main source. Note that by construction this statistic is larger than 50%, i.e. $\alpha_k > 1/2$. To get a visual sense of the degree of disagreement in the intensive margin, Figure 4 plots the distribution of expenditure shares for the products with the average, median, 25th and 75th percentile levels of dispersion, as measured by their σ_k. Expenditure shares are demeaned within each product-year and pooled across years. Perfect agreement would correspond to a degenerate histogram, with all the mass at zero. In contrast,

[44]Note that in the previous section we "rejected" the weaker prediction embedded (8), so that (7) is also rejected.

[45]That is, $\alpha_k = \sum 1(s_{M(k),i} > 0.5)/N_k$, where N_k is the number of firms in the subsample corresponding to product k.

[46]This corresponds to the standard deviation of the residual from a regression of $s_{M(k),i}$ on sector fixed effects.



Notes: Each graphs plots the distribution of expenditure shares on a particular variety for firms that agree in their extensive margin of trade. The products chosen are the ones that feature the average, median, 25th percentile and 75th percentile levels of σ_k, respectively. For every product-year, expenditure shares are demeaned, and then pooled across years within a product.

Figure 4: Disagreement on the Intensive Margin

we see that in all four products a substantial fraction of the firms are located away from zero. Even for the product with the least dispersion (bottom left panel), more than 70% of the firms feature an expenditure share that is different from the average.

The disagreement documented in Table 6 and Figure 4 suggests that - through the lens of our baseline model - price adjusted qualities exhibit considerable variation across firms. Furthermore, the results of the previous section indicate that this disagreement is not the consequence of only firm-specific idiosyncratic noise in quality adjusted prices, but that expenditure shares are systematically related to firm productivity.

Finally, we should note that the results of this section have an important methodological implication. In principle, we could have run the regressions of the previous section holding the identity of a particular variety fixed - i.e. not using order statistics. Interestingly, when we do so in our data, we find $\beta \approx 0$, i.e. we would find support for the homothetic demand system outlined in Section 2.[47] As shown above, this conclusion would of course be erroneous. The reason turns out to be precisely the disagreement found in this section. To see this, consider the following simple example. Consider a group of firms that source a given product from countries A and B and suppose there are two types of firms: high and low productivity. Suppose that firms' intensive margin behavior is indeed non-homothetic in the following way: more productive firms spend more on their most important variety. For example, high productivity firms spend 80% of their budget in their most important country, while low productivity ones spend only 60%. Crucially, firms disagree on the identity of

[47]Regression results are available from the authors upon request.

their most important supplier: within each productivity type, only 50% of the firms consider A as their main country. This could reflect the fact that, for these firms, country A offers superior quality (or lower price), while for the other firms the reverse is true. Under this conditions, a regression of expenditure shares on country A on productivity would yield a coefficient for productivity of zero: both productivity types feature an average expenditure share on country A of 50%. In contrast, a regression of the max share on productivity would correctly pick up the non-homotheticity. Thus, when firms exhibit disagreement in their ranking of varieties, using order statistics of expenditure shares to test for non-homotheticities is preferable to keeping identities fixed.

4 Mechanisms of Non-Homothetic Demand

As seen above, firms' intensive margin of import demand is non-homothetic as bigger firms concentrate their spending on their most important supplier. Importantly, this conclusion does not depend on how the extensive margin of trade is determined as all results explicitly control for firms' sourcing strategy. Hence, any model that is characterized by (i) factor neutral productivity and constant returns to scale and (ii) common input prices across firms[48] will not be able to account for this moment of the micro data. A microfounded theory of import behavior therefore has to contain a mechanism that can generate non-homothetic import demand in the cross-section of firms.

In this section, we briefly discuss three mechanisms that can generate such non-homotheticity. In particular, we show that our results are consistent with a complementarity between input quality and firm productivity as in Kugler and Verhoogen (2011), a simple process of search where bigger firms sample more supplying firms *within* sourcing countries, or the existence of unobserved intra-firm trade. As will be clear below, we cannot distinguish between these mechanisms with our data, which only covers the demand side of import transactions.[49] Hence, we think of these as plausible microfoundations that can be added to a richer theoretical framework to quantitatively account for the micro-data.

Quality-Productivity Complementarity To see that the non-homotheticity documented in Section 3 is consistent with the presence of a complementarity between input quality and firm productivity, consider the following tractable extension of the basic framework in the spirit of Kugler and Verhoogen (2011). We continue to assume that production features constant returns to scale, that

[48] We require that, conditional on importing a particular variety, all firms face the same price.

[49] However, we want to stress that if one is only interested to introduce an additional degree of freedom in the theory to be able to account for the empirical evidence, one might not even be interested in the true data generating process. As an analogy, consider the case of exports. In the data, we see firms exporting small amounts. At the same time, few firms are exporters. A basic Melitz (2003) model has problems to match this fact as the level of fixed costs determine both the productivity of the marginal exporter and the number of exporting firms. To improve the empirical fit along this dimension, Eaton, Kortum, and Kramarz (2011) borrow the insight of Arkolakis (2010) and introduce a marketing technology, where firms' extensive margin of sales (within countries) is increasing in firm productivity. While this microfoundation seems very plausible (and Arkolakis (2010) presents empirical evidence), it is arguably less important whether this is the actual mechanism, as long as the focus is mostly on developing a quantitative framework to match the microdata for French exporters.

productivity is the single relevant source of heterogeneity[50] and that firms face common prices after entry. However, we drop the assumption that productivity is purely factor-neutral and impose more structure on how productivity (a firm-characteristic) and variety-quality (a country characteristic) interact to generate firm-specific quality-flows. Letting η^i_{ck} be firm i's quality flow of variety c for product k, we write $\eta^i_{ck} = \eta_k(q_{ck}, \varphi_i)$, where q_{ck} denotes the quality of a variety c in product k and η_k is a product-specific function.[51] A natural question is whether productivity φ and quality q are complements or substitutes in production: do productive firms have a bias towards high-quality varieties or do they have a comparative advantage in low-quality inputs, because their productivity can substitute for the inferior innate quality of sourced products? While in principle both substitutability and complementarity are plausible[52], we now argue that our findings in Section 3.3 suggest that productivity and quality are complements. To see this, consider the CES model from the previous section. From (11), the relative expenditure shares between two varieties c and c' for product k are given by

$$ln\left(\frac{s_{ck}(\varphi,\Sigma)}{s_{c'k}(\varphi,\Sigma)}\right) = (\rho-1)\,ln\left(\frac{\eta_k(q_{ck},p_{ck},\varphi)}{\eta_k(q_{c'k},p_{c'k},\varphi)}\right) - (\rho-1)\,ln(\frac{p_{ck}}{p_{c'k}}). \tag{18}$$

Note that while the within-firm ranking of different sourcing countries is still determined by country-characteristics (q_{ck}, p_{ck}), relative expenditure shares are now dependent on firm-productivity.[53]

Now consider the results in Tables 4 and 5, where we regressed relative expenditure shares on sales - see equation (17). While the theory in Section 3.3 solely implied that the coefficient for sales should be zero, (18) now shows that the sign of sales is informative: holding prices fixed, expenditure shares are increasing in quality. Hence, under fixed prices, our finding that bigger firms bias their expenditure towards high quality inputs implies that η is log supermodular, i.e. that quality and productivity are complements. However, without more information on how prices prices and qualities co-move, it is not possible to infer the relative ranking of qualities from the distribution of expenditure.[54]

To make progress, we exploit the information contained in unit values, which we observe directly in the data and refer to as prices. By observing the price of different varieties, we can directly study the log-supermodularity of η by focusing on firms' *average prices.* In particular, define firm i's

[50] As noted before, we also allow for unrestricted heterogeneity in firms' fixed costs. This affects the extensive margin of trade but not the allocation of expenditure across sourcing countries.

[51] Note that this formulation is simply the limiting the case of introducing a second source of firm-heterogeneity A, which determines quality flows by $\eta^i_{ck} = \eta_k(q_{ck}, A_i)$. Hence, firms are characterized by the vector (φ, A), where φ is factor-neutral productivity and A parametrizes the firm-specific "quality-bias". Our formulation is the case where φ and A are perfectly correlated.

[52] If we interpret firm productivity φ as having a quality component (i.e. bigger firms are firms producing high quality goods), we would expect firm productivity and input quality as being complements - in order to produce high quality output, high quality inputs are required. If on the other hand we think of $\eta(.)$ as being a reduced form for the quality of the interaction between foreign suppliers and the importing firm, it is reasonable to think of productivity and input quality as substitutes, if for example superior managerial skills at the firm level can better deal with low-quality inputs like misspecified intermediate products or delays in the supply chain.

[53] In fact, it is not only the case that different firms have different expenditure shares for individual varieties, but different firms do not necessarily agree on the ranking of different countries. This pattern of imperfect agreement is in line with the evidence reported in Section 3.4 above.

[54] Concretely, a firm may spend a higher share of its budget for product k on Country 1 vs Country 2 even if $q_{2,k} > q_{1,k}$, as long as $p_{1,k}$ is sufficiently lower than $p_{2,k}$.

average price for product k given a sourcing strategy Σ_k as

$$\omega_k^i(\Sigma_k) = \sum_{c \in \Sigma_k} p_{ck} s_{ck}^i = \sum_{c \in \Sigma_k} p_{ck} \frac{p_{ck} z_{ck}^i}{\sum_{j=1}^{C_k} p_{jk} z_{jk}^i}. \tag{19}$$

Then, as shown in the Appendix, $\omega_k^i(\Sigma_k)$ is increasing in firm-productivity φ_i if and only if (i) p_{ck} is increasing in q_{ck} and η is log-supermodular or (ii) p_{ck} is decreasing in q_{ck} and η is log-submodular. This result is very intuitive. Firms pay high average prices for a given product k if they spend a relatively large amount of their budget on expensive varieties. If prices and qualities are positively related and more productive firms pay higher average prices, it follows that higher productivity induces a bias towards high quality varieties, which amounts to a log-supermodular η. If low-quality inputs were more expensive, a positive correlation between average prices ω_k^i and firm-productivity would be indicative of productivity and quality being substitutes. We follow Kugler and Verhoogen (2011) and the large literature in international trade that argues that high quality products are more expensive.[55]

To test this prediction, we again need to make sure that the variation in average prices is entirely an intensive margin phenomenon. In the most parsimonious specification we follow the approach from above and control for the sourcing strategy simply by its cardinality, i.e. the number of varieties sourced. Hence, consider first the top panel of Table 7, where we calculate ω_k^i according to (19) and then run the regression

$$ln\left(\omega_k^i\right) = \alpha_k + \phi ln\left(\left|\Sigma_k^i\right|\right) + \beta ln\left(S_i\right) + X_i'\gamma + u_{ck}^i, \tag{20}$$

where α_k is a product-fixed effect, $\left|\Sigma_k^i\right|$ is the number of varieties sourced, S_i denotes firm sales and X_i is a set of firm-specific controls. Column 3 clearly shows that β is robustly positive: for a given sourcing strategy Σ_k, larger firms consistently pay more for their product-specific import bundle. If more expensive inputs are of higher quality, this is evidence of a complementarity between firm productivity and input quality. Quantitatively, these results imply that a one standard deviation increase in sales increases the average price paid for a given product bundle by 13%. Column 4 shows that this point estimate hardly changes after we control for other firm characteristics.

Columns 1 and 2 show an interesting asymmetry between the intensive and extensive margin of importing. For the case of exports, we know that more productive firms are more likely to sell in poor countries. Eaton, Kortum, and Kramarz (2011, p. 1453) for example report that "average sales

[55]Schott (2004) for example exploits aggregate data from US customs forms to shows that richer countries have consistently higher unit values and concludes that "high-wage countries export vertically superior varieties". Similarly, Hummels and Klenow (2005, p. 705) conclude that "richer countries export higher quantities of each good at modestly higher prices, consistent with higher quality" and similar results are obtained in Khandelwal (2010) and Hallak and Schott (2011). We test this relationship in our French dataset. Specifically, we run a regression of the form $ln\left(\overline{p}_{ck}\right) = \delta_k + \alpha ln\left(y_c\right) + u_{ck}$,where $\overline{p}_{ck}$ is the average unit value of product k imported from country c (where the average is taken across firms), y_c denotes income per capita of country c and δ_k is a product fixed effect. In our sample, we estimate $\alpha = 0.14$, which is highly significant and almost numerically identical to the one reported in Schott (2004). In a rare exception, where quality can be "objectively" assessed, Crozet, Head, and Mayer (2011) study export prices for French champagne and show that champagne of higher quality consistently sells at higher prices.

	Dep. variable: log Average price ($\omega_k^i = \sum_{c=1}^{C_{k,i}} p_{ck}^i s_{ck}^i$)			
Panel A	Controlling for the number of countries sourced			
ln sales	0.064***		0.067***	0.064***
	(0.001)		(0.001)	(0.001)
ln nb. varieties (prod.)		-0.021***	-0.077***	-0.072***
		(0.004)	(0.004)	(0.004)
Firm Controls	No	No	No	Yes
N	610,375	610,375	610,375	552,722
R^2	0.74	0.74	0.74	0.74
Panel B	Controlling for the sourcing strategy			
ln sales	0.048***	0.045***		
	(0.002)	(0.003)		
Firm Controls	No	Yes		
N	610375	552722		
R^2	0.92	0.92		

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. Specification with firm controls control for (log) capital intensity and export status and contain indicator variables if the firm is member of a foreign or a corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment. We consider all firm-product pairs, who source at least two varieties of the respective product. In Panel A we control for the sourcing strategy by the number of sourcing countries. In Panel B we control for the sourcing strategy by sourcing strategy specific fixed effects.

Table 7: Firm Size and Average Import Prices

in France rise systematically with selling to less popular markets" as an important regularity in the micro data. This is a statement about the extensive margin of exporting. In our regressions, this extensive margin of importing is captured by the coefficient ϕ, which is also negative. Hence, as in the exporting literature, the quality of the *marginal* sourcing country is in fact decreasing, i.e. firms that source a particular product from more countries (i.e. bigger firms) source on average a lower quality. However, on the intensive margin, this cross-sectional correlation is reversed: more productive firms have a bias towards high-quality countries and it is this intensive margin that dominates the simple bivariate correlation between prices and sales as seen in the first column.

In Panel B of Table 7 we then follow our approach of Table 4 to correctly control for firms' extensive margin by introducing sourcing-strategy dummies. Once again, we find robust evidence that more productive firms pay more for their inputs. While the point estimates drop by 25%, they remain highly significant.

The results of Table 7 are in line with those in Kugler and Verhoogen (2011), who use Colombian data to document that larger plants pay more for their inputs.[56] There are, however, two differences. First, we focus only on imported inputs while they also consider domestic inputs as well. Second, we explicitly control for the extensive margin of sourcing while they document the unconditional relationship between inputs prices and plant size. Thus, we are able to more clearly distinguish between technological complementarities and non-homotheticities induced by fixed costs.

Search Instead of directly introducing the non-homotheticity in firms' production technology, we can also interpret it as reflecting unobserved differences in firms' choice sets stemming from search frictions.[57] Suppose that productivity differences are actually factor-neutral, but that there are many producers *within* a sourcing country that can produce a particular variety. Suppliers are heterogeneous in the (price-adjusted) quality ξ they offer and firms search for a supplier by taking draws from a (variety specific) distribution $G_{ck}(\xi)$. If n suppliers are sampled, the actual quality firm i buys in equilibrium is the maximum $\xi_{ck}^i = \max\left\{\xi_{ck}^j\right\}_{j=1}^{n_i}$, i.e. equilibrium qualities are firm-specific through a process of selection. If search is costly, more productive firms will in general search more intensely, because they will buy larger quantities. As the distribution of ξ_{ck}^i is increasing in n_i (in a first-order stochastic dominance sense), bigger firms naturally will have better qualities (on average) within countries. Note that this is not sufficient to generate the non-homotheticity we find: if bigger firms receive better offers in *all* their supplying countries, the allocation of expenditure within firms across countries will *not* necessarily be different across producers. For a process of search to generate the pattern we find in the data, it has to be the case that the distribution of the order statistics of ξ_{ck}^i *across* countries is increasing (in a first-order stochastic dominance sense) in the intensity of search. Hence, in the same way as the production function for firm-specific qualities $\eta_k(q_{ck}, \varphi_i)$ above has to satisfy a restriction for the second derivative, the underlying process of search has to generate a pattern where more productive firms are more likely to find *relatively* better suppliers

[56] Kugler and Verhoogen (2011) proxy productivity by employment while we use sales. We re did their analysis in our data using employment as a measure of productivity and find similar results.

[57] See e.g. Allen (2012) for an application of search frictions in international trade.

across countries. Whether that is actually the case depends heavily on the underlying distribution of qualities $G_{ck}(\xi)$. In Section 6.7 in the Appendix we construct two explicit examples where search either generates or does not generate the cross-sectional pattern of spending we see in the data. Hence, while we think that experimentation in the market for suppliers offers a plausible mechanism for our results, it does not automatically generate the pattern we see in the data.

Intra-firm trade As a third example, consider the occurrence of intra-firm trade. As in Melitz, Helpman, and Yeaple (2004), suppose French importers have two choices how to source an input from abroad - they can either import it on the spot market or they can directly invest into foreign production facilities (FDI). While importing directly has low fixed and high variable costs, owning the supplier directly has high fixed and low variable costs. As firms' expenditure shares are determined by price-adjusted qualities, a reduction in variable trade costs is isomorphic - from the point of view of the firm - to an increase in quality.

Now suppose that small importers only use spot market trading as their small import volume does not justify the expense of the fixed costs. Similarly, suppose that large importers do engage in FDI in some of their products. As FDI is subject to increasing returns (due to the fixed costs), large importers will direct their FDI activities to trading partners that account for a large part of their import budget. Hence, large firms will perceive relatively high quality flows for products that are relatively important - which is exactly the non-homotheticity we document in the data.

Indeed we have some information in our data to try to partially address this concern. In our firm-level data we observe if a French firm has a foreign affiliate. We do not observe this information by product. However, we redid our analysis after dropping all firms with such affiliations. The results are contained in Table 12 in the Appendix. The point estimates are almost identical to our main specifications. In particular, we still find robust evidence for the non-homotheticity of import demand in the sample of firms who report to not have any foreign affiliates. While this should address the most pressing concerns about intrafirm-trade, we want to stress that the legal definition of foreign affiliations (which is the one we observe the data) might be a very noisy measure for this phenomenon. In general, the same pattern could emerge from any quality-increasing relationships between foreign suppliers and French importers irrespective of their legal status. If more productive firms are more likely to engage in such relationships, the data will look as if the demand-system was non-homothetic.

5 Conclusion

In this paper we use micro-data to study the import behavior of French manufacturing firms. We look at the data through the lens of a static model of importing that features complementarities across imported inputs and variety-specific fixed costs to access international markets. The model has two main ingredients. We assume that (i) productivity is factor neutral and production is subject to constant returns and (ii) firms face the same international prices. Other than that, we allow for unrestricted firm heterogeneity in both productivity of the fixed costs of importing. Hence, this

framework is not only general enough to nest the available models of import demand as special cases, but it arguably also represents the backbone of a framework of import demand that can be eventually estimated with firm-level data.

Our first result is theoretical and concerns the empirical content of this framework. In contrast to the case of exporting, there are no robust predictions on the extensive margin for imports. In particular, the theory does not imply a sorting condition by which more productive firms access more international markets. This result is not so much due to particular modeling assumptions, but follows from the inherent asymmetry between import and exports: while the decision to export can be made "market-by-market" (Eaton, Kortum, and Kramarz, 2011), import decisions are necessarily interdependent as long as the production function features love-for-variety, i.e. imported inputs are not perfect substitutes.

In contrast, we show that the theory has a robust prediction for the intensive margin of trade: conditional on the sourcing strategy, i.e. firms' extensive margin of trade, expenditure shares across products and varieties are fully determined by price-adjusted qualities, that is, by characteristics of the supplying country. In particular, firm productivity should not affect relative input demand once the sourcing strategy is controlled for. This exclusion restriction is independent on how the extensive margin is determined and hence is present in all models of importing that we are aware of. We show that this homotheticity property is not supported by the data. Moreover, the particular direction in which the theory is rejected is economically meaningful: holding firms' sourcing strategy fixed, more productive firms spend a higher share of their budget on their most important supplier. Hence, the data asks for an additional mechanism for why big importers concentrate their spending on their top trading partners. We discuss three plausible possibilities. First of all, this pattern of non-homothetic demand is consistent with a model where firm productivity and input quality are complements. Secondly, it could be the outcome of a process of search whereby French importers sample different suppliers within foreign countries and bigger firms search more intensely. Finally, it could be a consequence of particular buyer-supplier relationships akin to theories of intra-firm trade.

References

Allen, T. (2012): "Information Frictions in Trade," Working Paper.

Amiti, M., O. Itskhoki, and J. Konings (2012): "Importers, Exporters, and Exchange Rate Disconnect," Working Paper.

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

Arkolakis, C. (2010): "Market Penetration Costs and the New Consumer Margin in International Trade," *Journal of Political Economy*, 118(6), 1151–1199.

Arkolakis, C., and M.-A. Muendler (2011): "The Extensive Margin of Exporting Products: A Firm-level Analysis," Working Paper.

Armenter, R., and M. Koren (2008): "A Balls-and-Bins Model of Trade," Working Paper.

Bernard, A., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," *American Economic Review*, 93(4), 1268–1290.

Bernard, A. B., and J. B. Jensen (1999): "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Economics*, 47(1), 1–25.

Bernard, A. B., J. B. Jensen, S. Redding, and P. Schott (2007): "Firms in International Trade," *Journal of Economic Perspectives*, 3(21), 105–130.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2012): "The Empirics of Firm Heterogeneity and International Trade," *Annual Review of Economics*, 4, 283–313.

Bernard, A. B., S. Redding, and P. Schott (2010): "Multi-Product Firms and Product Switching," *American Economic Review*, 100(1), 70–97.

Bernard, A. B., S. J. Redding, and P. K. Schott (2011): "Multi-Product Firms and Trade Liberalization," *Quarterly Journal of Economics*, 126(3), 1271–1318.

Broda, C., and D. E. Weinstein (2006): "Globalization and the Gains from Varieties," *Quarterly Journal of Economics*, 121(2), 541–585.

Chaney, T. (2008): "Distorted Gravity: the Intensive and Extensive Margins of International Trade.," *The American Economic Review*, 98(4), 1707–1721.

——— (2013): "The Network Structure of International Trade," .

Crozet, M., K. Head, and T. Mayer (2011): "Quality Sorting and Trade: Firm-level Evidence for French Wine," *Review of Economic Studies*, 79(2), 609–644.

Eaton, J., and S. Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70(5), 1741–1779.

Eaton, J., S. Kortum, and F. Kramarz (2004): "Dissecting Trade: Firms, Industries, and Export Destinations," *The American Economic Review*, Vol. 94(2), 150–154.

——— (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79(5), 1453–1498.

Foster, J., J. C. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Goldberg, P., A. Khandelwal, N. Pavcnik, and P. Topalova (2010): "Imported Intermediate Inputs and Domestic Product Growth: Evidence from India," *Quarterly Journal of Economics*.

Gopinath, G., and B. Neiman (2012): "Trade Adjustment and Productivity in Large Crises," Working Paper.

Grossman, G. M., E. Helpman, and P. Kircher (2013): "Matching and Sorting in a Global Economy," .

Hallak, J. C., and P. K. Schott (2011): "Estimating Cross-Country Differences in Product Quality," *Quarterly Journal of Economics*, 126, 417–474.

Halpern, L., M. Koren, and A. Szeidl (2009): "Imported Inputs and Productivity," Working Paper.

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

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

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

——— (2011): "Prices, Plant Size, and Product Quality," *Review of Economic Studies*, (14418), 1–33.

Lapham, B., and H. Kasahara (2013): "Productivity and the Decision to Import and Export: Theory and Evidence," *Journal of International Economics*, 89(2), 297–316.

Levinsohn, J., and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–341.

Manova, K., and Z. Zhang (2012): "Export Prices Across Firms and Destinations," *Quarterly Journal of Economics*, 127, 379–436.

Mayer, T., M. Melitz, and G. Ottaviano (2010): "Market size, Competition, and the Product Mix of Exporters," Working Paper.

Melitz, M. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

Melitz, M., E. Helpman, and S. Yeaple (2004): "Export Versus FDI with Heterogeneous Firms," *The American Economic Review*, 94(1), 300–316.

Melitz, M. J., and S. J. Redding (2012): "Heterogeneous Firms and Trade," NBER Working Paper 18652.

Ramanarayanan, A. (2012): "Imported Inputs and the Gains from Trade," Working Paper.

Redding, S. (2011): "Theories of Heterogeneous Firms and Trade," *Annual Review of Economics*, 3, 77–105.

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

Tintelnot, F. (2012): "Global Production with Export Platforms," Working Paper.

6 Appendix

6.1 Data Description

Our main data set stems from the information system of the French custom administration (DGDDI) and contains the universe of import and export flows by French manufacturing firms. The data is collected at the 8-digit (NC8) level and a firm located within the French metropolitan territory must report this detailed information as long as the following criteria are met. Within EU imports, have to be reported as long as the firm's annual trade value exceeds 150,000 Euros. If that threshold is not met, firms can choose to report under a simplified scheme. However, in practice, many firms under that threshold report the detailed information. For imports from outside the EU, all shipments must be reported to the custom administration as this data is used to calculate the value added tax in all cases. The conditions are more stringent for exports. For within EU exports, all shipments must be reported to the custom administration. For exports outside the EU, reporting is required if the exported value to exceeds 1,000 Euros or weighs more than a ton.

The attractive feature of the French data is the presence of unique firm identifiers (the SIREN code), which is available in all French administrative files. Hence, various other datasets can be matched to the trade data at the firm level. To learn about the characteristics of the firms in our sample we employ fiscal files.[58] Sales are deflated using price indices of value added at the 3 digit level obtained from the French national accounts. To measure the expenditure on domestic inputs, we subtract the total import value from the total expenditure on wares and inputs reported in the fiscal files. Capital, used for the TFP estimation, was computed using a permanent inventory method. The series were initialized with the deflated value of assets reported in the first year of reported fiscal account (1995). We then used the reported investment expenditure, which we deflated with an investment price index available from the French national accounts. We assumed a depreciation rate of 10%.

Additionally we use the French business registers (SIRENE files), created by the Firm Demography Department of the French National Institute of Statistics (INSEE). The SIRENE files report the yearly creation and destruction of French firms and provide us with information about firm age and

[58]The firm level accounting information is retrieved from two different files: the BRN ("Bénéfices Réels Normaux") and the RSI ("Régime Simplifié d'Imposition"). The BRN contains the balance sheet of all firms in the traded sectors with sales above 730,000 Euros. The RSI is the counterpart of the BRN for firms with sales below 730,000 Euros. Although the details of the reporting differs, for our purpose these two data sets contain essentially the same information. Their union covers nearly the entire universe of all French firms.

	Full Sample	Importers	Non-Importers	Exporters	Non-Exporters
Employment	18.18	89.72	5.71	78.89	5.9
Sales	4342.35	23646.8	969.15	20842	1035.52
Sales / Worker	104.42	157.62	92.3	150.46	92.3
Capital / Worker	33.74	44.28	31.29	42.45	31.38
Inv. / Worker	2.9	4.13	2.6	3.91	2.61
Inputs (mat.)	0.2	0.3	0.18	0.28	0.18
Import share	0.04	0.3	0	0.2	0.01
Share of Importers	0.15	1	0	0.66	0.05
Share of Exporters	0.17	0.75	0.07	1	0
Firm Age	14.38	19.45	13.51	19.54	13.36
Foreign owned	0.02	0.13	0.01	0.11	0.01
Foreign Group	0.05	0.23	0.01	0.2	0.02
Labor Productivity	43.783	56.496	40.726	56.023	40.405
TFP-LP	28.135	30.255	27.788	29.23	27.999
TFP-OLS	25.412	26.519	25.251	26.61	25.189
Number of Firms	259,602	31,022	228,580	34,527	225,075

Notes: Sales, wages, expenditures on imports or exports are all expressed in 2000 prices using a 3-digit industry level price deflator. Our capital measure is the book value reported in firms' balance sheets ("historical cost"). We measure employees by occupation. Skilled workers are engineers, technicians and managers, workers of intermediate skills are skilled blue and white collars and low skilled workers are members of unskilled occupations. A firm is foreign owned, if the controlling entity is a foreign company. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France.

Table 8: Characteristics of importers, exporters and domestic firms

legal status. 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) describing the structure of ownership of all of the French firms in the private sector whose financial investments in other firms (participation) are higher than 1.2 million Euros or having sales above 60 million Euros or 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 with the register of firms that are controlled by the State.

Using these datasets, we construct a non-balanced panel dataset spanning the period from 2001 to 2006. Some basic characteristics of importing and non-importing firms are contained in Table 8. For comparison, we also report the results for exporting firms. Expectedly, importers outperform domestic firm in essentially all dimensions we look at (see also Bernard, Jensen, Redding, and Schott (2012)). Furthermore, import and export status are highly correlated.

6.2 Cross-Sectional Variation of the Number of Varieties Sourced

In Table 9 below we depict the distribution of the number of varieties sourced per product within narrowly defined industries. While there are some cross-industry differences, it is clearly seen that the heterogeneity in the number of varieties sourced is mostly a within-industry-across-firm phenomenon and not only driven by sectoral differences.

	Mean	Percentiles						No of firms	Agg. Import share
		25	50	75	90	95	99		
Food Products	1.803	1	1.179	1.960	3.042	4.145	9.354	3065	0.074
Clothing, Leather	1.817	1	1.275	2.083	3.282	4.224	7.586	1900	0.028
Editing and Printing	1.800	1	1	1.997	3.516	4.846	8.510	1794	0.011
Pharmaceuticals	2.085	1	1.576	2.397	3.823	5.090	8.886	714	0.073
Furnitures etc	1.942	1	1.233	2.180	3.706	5.012	8.650	2414	0.044
Car Industry	2.615	1.006	1.890	3.373	5.495	7.040	10.088	613	0.060
Other Transport Equip.	2.524	1	1.218	2.549	5.182	8.549	17.329	535	0.191
Mechanical Equip.	2.002	1	1.242	2.150	3.872	5.344	9.816	3568	0.047
Elec. and Electron. Equip.	2.362	1	1.247	2.621	4.870	6.922	12.807	1719	0.057
Mineral Products	1.704	1	1.093	1.983	3.006	4.036	6.675	1269	0.015
Textile	1.966	1	1.529	2.403	3.604	4.628	7.017	1484	0.018
Paper and Wood	1.817	1	1.268	2.076	3.249	4.172	6.745	1982	0.030
Chemical, Rubber, Plastic	2.111	1	1.642	2.537	3.883	5.040	8.625	2833	0.092
Metal Work	1.752	1	1.071	1.993	3.207	4.316	7.045	3541	0.053
Elec. and Electron. Compo.	2.346	1	1.618	2.841	4.619	6.281	10.844	878	0.033
Coke, Ref. Petrol., Nuclear	2.679	1.005	1.718	2.871	4.505	9.329	17.203	69	0.174
Total Economy	1.961	1	1.281	2.173	3.700	4.962	8.938	28378	1

Notes: The table displays sector-specific properties of the distribution of firms' average number of varieties per product, i.e. $C_i = \sum_k s_{k,i} C_{k,i}$, where $s_{k,i}$ is firm i's expenditure share on product k and $C_{k,i}$ is the number of countries firm i sources product k from. The final columns contain the number of firms in the respective sector and the share of French imports the respective sector accounts for.

Table 9: Firm-heterogeneity in variety sourcing within industries

6.3 Optimal Import Demand and the Exclusion Restriction

Consider the minimization problem (5), i.e. $\min_z \left\{ \sum_{(c,k)\in\Sigma} p_{ck} z_{ck} \text{ s.t. } q(z) \geq \frac{y}{\varphi} \right\}$. Let us consider the problem in efficiency units $\tilde{z}_{ck} \equiv \eta_{ck} z_{ck}$with a price $\tilde{p}_{ck} \equiv \frac{p_{ck}}{\eta_{ck}}$. The first order conditions are given by

$$\tilde{p}_{ck} = \lambda \frac{\partial f(x)}{\partial x_k} \frac{\partial g_k(\tilde{z})}{\partial \tilde{z}_{ck}}, \tag{21}$$

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

$$\sum_{c\in\Sigma_k} \tilde{p}_{ck} \tilde{z}_{ck} = \lambda \frac{\partial f(x)}{\partial x_k} \sum_{c\in\Sigma_k} \frac{\partial g_k(\tilde{z})}{\partial \tilde{z}_{ck}} \tilde{z}_{ck} = \lambda \frac{\partial f(x)}{\partial x_k} x_k,$$

where the last equality follows from g_k being CRS.[59] Hence,

$$\Gamma(\Sigma, y, \varphi) = \sum_{k\in K} \sum_{c\in\Sigma_k} \tilde{p}_{ck} \tilde{z}_{ck} = \lambda \sum_{k\in K} \frac{\partial f(x)}{\partial x_k} x_k = \lambda q(\tilde{z}) = \lambda \frac{1}{\varphi} y, \tag{22}$$

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

$$\sum_c \frac{\partial g_k(\mu\tilde{z}_{1k}, ..., \mu\tilde{z}_{Ck})}{\partial(\mu\tilde{z}_{ck})} \tilde{z}_{ck} = g_k(\tilde{z}_{1k}, ..., \tilde{z}_{Ck}).$$

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

so that $\frac{\lambda}{\varphi}$ are the marginal cost of production. The expenditure share of variety 1 in product k is given by

$$s_{1k} = \frac{\tilde{z}_{1k}\tilde{p}_{1k}}{\sum_{c\in\Sigma_k}\tilde{p}_{ck}\tilde{z}_{ck}} = \frac{\frac{\partial g_k(\tilde{z})}{\partial \tilde{z}_{1k}}\tilde{z}_{1k}}{g_k\left(\tilde{z}_{1k},\tilde{z}_{2k},...,\tilde{z}_{nk}\right)} = \frac{\frac{\partial}{\partial \tilde{z}_{1k}}g_k\left(1,\chi_{2k},...,\chi_{nk}\right)}{g_k\left(1,\chi_{2k},...,\chi_{nk}\right)},$$

where $\chi_{ck} \equiv \frac{\tilde{z}_{ck}}{\tilde{z}_{1k}}$. We have to show that χ_{ck} depends only on the set of prices $[\tilde{p}_{ck}]_c$. (21) implies that

$$\left[\lambda\frac{\partial f(x)}{\partial x_k}\right]^{-1} = \frac{\partial g_k(\tilde{z})}{\partial \tilde{z}_{c'k}}\frac{1}{\tilde{p}_{c'k}} = \frac{\partial g_k(\tilde{z})}{\partial \tilde{z}_{ck}}\frac{1}{\tilde{p}_{ck}} \text{ for all } c, c' \in C_k. \tag{23}$$

As g_k is constant returns to scale, (23) implies that

$$\frac{\partial g_k(1,\chi_2,..,\chi_n)}{\partial \tilde{z}_{c'k}} = \frac{\partial g_k(1,\chi_2,..,\chi_n)}{\partial \tilde{z}_{1k}}\frac{\tilde{p}_{ck}}{\tilde{p}_{1k}} \text{ for all } c \in C_k.$$

These are $|C_k|-1$ equation in $|C_k|-1$ unknowns $(\chi_2,...,\chi_n)$ which have a solution $\chi_j = \rho_j\left([\tilde{p}_{ck}]_c\right)$ for all $j \neq 1$. Hence, as required we have that

$$s_{ck}(\Sigma, y, \varphi) = h_k\left([\tilde{p}_{ck}]_c\right).$$

To derive the cost function, note that (22) and (23) imply that[60]

$$\Gamma(\Sigma, y, \varphi) = \lambda\frac{1}{\varphi}y = \gamma\left(\{\tilde{p}_{ck}\}_{c,k}\right)\frac{1}{\varphi}y.$$

6.4 Extensive Margin of Importing

In Table 10 we report the results on the extensive margin of importing. In particular, we regress an indicator of import status (columns 1 and 2), the number of products sourced (columns 3 and 4) and the number of varieties sourced (columns 5 and 6) on different firm characteristics. It is clearly seen that empirically, firm productivity (as proxied for by sales) is strongly positively related to all these extensive margin measures. We do not focus on the extensive margin as the theory has less robust predictions for this margin of the data, but we report these results for completeness with the findings reported in the literature (Halpern, Koren, and Szeidl, 2009; Bernard, Jensen, Redding, and Schott, 2012, 2007).

[60]Note that $x_k = \tilde{z}_{1k}g_k(1,\chi_{1k},...,\chi_{Ck}) = \tilde{z}_{1k}\phi_k\left([\tilde{p}_{ck}]_c\right)$ so that (23) implies for all $k = 2,...,K$ that

$$\begin{aligned}\frac{\tilde{p}_{1k}}{\tilde{p}_{11}} &= \frac{f_k\left(\tilde{z}_{11}\phi_1\left([\tilde{p}_{c1}]_c\right),..,\tilde{z}_{1K}\phi_K\left([\tilde{p}_{cK}]_c\right)\right)}{f_1\left(\tilde{z}_{11}\phi_1\left([\tilde{p}_{c1}]_c\right),..,\tilde{z}_{1K}\phi_K\left([\tilde{p}_{cK}]_c\right)\right)}\frac{\partial g_k(\tilde{z})/\partial \tilde{z}_{1k}}{\partial g_1(\tilde{z})/\partial \tilde{z}_{11}} \\ &= \frac{f_k\left(\phi_1\left([\tilde{p}_{c1}]_c\right),..,\eta_K\phi_K\left([\tilde{p}_{cK}]_c\right)\right)}{f_1\left(\phi_!\left([\tilde{p}_{c1}]_c\right),..,\eta_K\phi_K\left([\tilde{p}_{cK}]_c\right)\right)}\frac{\psi_k\left([\tilde{p}_{cK}]_c\right)}{\psi_1\left([\tilde{p}_{cK}]_c\right)},\end{aligned}$$

where $\eta_k = \frac{\tilde{z}_{1k}}{\tilde{z}_{11}}$. These $K-1$ equations determine $\eta_k = \alpha\left(\{\tilde{p}_{ck}\}_{c,k}\right)$. Hence,

$$\lambda^{-1} = \frac{\partial f\left(\phi_1\left([\tilde{p}_{c1}]_c\right),\eta_2\phi_2\left([\tilde{p}_{c2}]_c\right),..,\eta_K\phi_K\left([\tilde{p}_{cK}]_c\right)\right)}{\partial x_k}\frac{\partial g_k(\tilde{z})}{\partial \tilde{z}_{c'k}}\frac{1}{\tilde{p}_{c'k}} = \vartheta\left(\{\tilde{p}_{ck}\}_{c,k}\right).$$

	Import Status		log nb of products sourced		log nb of varieties sourced	
log Sales	0.107***	0.029***	0.493***	0.436***	0.169***	0.151***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Lagged Import Status		0.018***				
		(0.004)				
Exporter		0.088***		0.292***		0.090***
		(0.002)		(0.007)		(0.003)
R^2	0.396	0.875	0.449	0.463	0.294	0.304
Observations	1,107,962	858,456	167,733	167,711	167,733	167,711

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year, age and 4-digit industry fixed effects. In columns (1) and (2), the dependent variable is an indicator of the firm's import status. In columns (3) and (4), the dependent variable is the log of the number of imported products of firm i, i.e. $ln(K_i)$, where $K_i = \sum_k 1\left[V_{k,i} > 1\right]$ and $V_{k,i}$ is firm i's number of varieties of product k. In columns (5) and (6), the dependent variable is the log of the average number of varieties of firm i, i.e. $ln(V_i)$, where $V_i = \sum_k s_{k,i} V_{k,i}$ and $V_{k,i}$ is firm i's number of varieties of product k and $s_{k,i}$ is firm i's expenditure share on product k. "Lagged Import Status" is an indicator for the firm's import status in the previous year. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate.

Table 10: The Extensive Margin

6.5 CES Regressions

While the Approach 3 in Section 3.3 above is attractive in that it focuses precisely on firms' intensive margin problem, the number of observations drops substantially. This issue becomes especially pressing for the case of 5 varieties, where the number of observations falls to approximately 6,000 and we have roughly 4,500 product fixed effects. To address this concern, we exploit 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 production functions take the CES functional form, the model 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:

$$ln(s^i_{jk}/s^i_{j'k}) = ln\left(s^i_{jk}\right) - ln\left(s^i_{j'k}\right) = (\rho - 1)\left(ln\left(\xi_{jk}\right) - \left(\xi_{j'k}\right)\right), \tag{24}$$

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 endogenous effect of the sourcing strategy, $\phi\left(\Sigma^i_k\right)$.[61] While equation (24) requires stronger assumptions than the general expenditure share equation (7) of Section 2, it is very useful. The key advantage of (24) 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. Thus, this approach results in an increase in sample size. In particular, to make (24) operational, we adopt the

[61] Of course (24) is only defined for pairs of varieties, which are actually sourced. That is, expression (24) is valid for $c, c' \in \Sigma^i_k$, and not defined otherwise. This means that when testing for the exclusion restriction implicit in (24) for two particular varieties we need to restrict ourselves to the set of firms that source those two varieties. The advantage of the differencing approach, relative to the previous one, 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.

following procedure. 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 previous approach. We then rank these V countries for each firm and run the regression contained in (17) on the enlarged subsamples.[62] Table 11, which has the same structure as Table 5, 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 negative and highly significant. As in Table 5, we also recover the monotonicity of the coefficients and the point estimates are very similar in magnitude. Hence, productive firms concentrate their spending on their most preferred sourcing countries relative to smaller importers. This non-homotheticity of import demand violates the exclusion restriction embedded in the theory of the benchmark model in Section 2.

6.6 Complementarity and Average Prices

Using the CES production structure, (19) implies that

$$\omega_k(\varphi) = \sum_{c\in\Sigma_k} p_{ck}\frac{p_{ck}z_{ck}^i}{\sum_{c\in\Sigma_k} p_{ck}z_{jk}^i} = \sum_{c\in\Sigma_k} p_{ck}\left(\frac{\left(\frac{\eta_k(q_{ck},\varphi)}{p_{ck}}\right)^{\rho-1}}{\sum_{j=1}^{C_k}\left(\frac{\eta_k(q_{jk},\varphi)}{p_{jk}}\right)^{\rho-1}}\right) \equiv \sum_{c\in\Sigma_k} p_{ck}\pi_{ck}(\varphi).$$

Differentiating with respect to φ we get that

$$\begin{aligned}
\frac{\partial\omega_k(\varphi)}{\partial\varphi} &= (\rho-1)\left[\sum_{c\in\Sigma_k} p_{ck}\frac{\partial ln\left[\eta_k(q_{ck},\varphi)\right]}{\partial\varphi}\pi_{kc}(\varphi) - \sum_{c\in\Sigma_k} p_{ck}\pi_{kc}(\varphi)\sum_{c\in\Sigma_k}\frac{\partial ln\left[\eta_k(q_{jk},\varphi)\right]}{\partial\varphi}\pi_{kc}(\varphi)\right] \\
&= (\rho-1)\left[E_\pi\left[p_{ck}\frac{\partial ln\left[\eta_k(q_{ck},\varphi)\right]}{\partial\varphi}\right] - E_\pi\left[p_{ck}\right]E_\pi\left[\frac{\partial ln\left[\eta_k(q_{ck},\varphi)\right]}{\partial\varphi}\right]\right] \\
&= (\rho-1)\,Cov_\pi\left(p_{ck}, \frac{\partial ln\left[\eta_k(q_{ck},\varphi)\right]}{\partial\varphi}\right),
\end{aligned}$$

where $E_\pi[.]$ denotes the expectation operator with respect to the measure, i.e. $E_\pi[x_{ck}] = \sum_{c\in\Sigma_k} x_{ck}\pi_{ck}$. Hence, $\frac{\partial\omega_k(\varphi)}{\partial\varphi} > 0$ if prices are increasing in q and η is log-supermodular or prices are decreasing in q and η is log-submodular.

6.7 Two simple models of search

In this section, we show that while the process of search outlined on page 31 could generate the non-homotheticity we find in the data, it does not do so automatically. In particular, it depends on the underlying distributions $G(\xi)$ the price-adjusted qualities of suppliers are drawn from. For

[62] 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.

	$\ln(s_2/s_1)$		$\ln(s_3/s_1)$		$\ln(s_4/s_1)$		$\ln(s_5/s_1)$	
				Firms sourcing the same 2 varieties of product k				
ln Sales	-0.102***	-0.085***						
	(0.003)	(0.004)						
Firm Controls	No	Yes						
N	189,140	171,539						
R^2	0.15	0.15						
				Firms sourcing the same 3 varieties of product k				
ln Sales	-0.053***	-0.054***	-0.140***	-0.130***				
	(0.004)	(0.005)	(0.006)	(0.008)				
Firm Controls	No	Yes	No	Yes				
N	69,134	62,870	69,134	62,870				
R^2	0.27	0.27	0.30	0.30				
				Firms sourcing the same 4 varieties of product k				
ln Sales	-0.034***	-0.031***	-0.088***	-0.089***	-0.178***	-0.179***		
	(0.006)	(0.007)	(0.008)	(0.010)	(0.010)	(0.013)		
Firm Controls	No	Yes	No	Yes	No	Yes		
N	34,539	31,386	34,539	31,386	34,539	31,386		
R^2	0.40	0.40	0.42	0.42	0.42	0.43		
				Firms sourcing the same 5 varieties of product k				
ln Sales	-0.010	-0.017	-0.042***	-0.055***	-0.096***	-0.106***	-0.162***	-0.164***
	(0.008)	(0.010)	(0.011)	(0.013)	(0.013)	(0.016)	(0.016)	(0.020)
Firm Controls	No	Yes	No	Yes	No	Yes	No	Yes
N	20,046	18,310	20,046	18,310	20,046	18,310	20,046	18,310
R^2	0.53	0.53	0.53	0.54	0.53	0.54	0.53	0.53

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. We weigh observations such that each firm has an equal weight. The dependent variables are $ln\left(s_j^i/s_1^i\right)$, which is the expenditure share on firm i's jth variety relative to its first variety, where the different varieties are ranked by their expenditure shares. The table contains 4 sets of regressions, each corresponding to a different number of varieties in the sourcing strategy. In any regression, all firms e.g. source from the same 2 varieties but might also source it from additional varieties (see main text for description of this procedure). All regressions that contain firm controls, control for firms' export status and (log) capital intensity and contain indicator variables if the firm is member of a foreign or a corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment.

Table 11: Firm Characteristics and the Intensive Margin of Trade: Exploiting the CES Structure

simplicity, consider the case of firms sourcing from exactly two countries ($c = 1, 2$), suppose that the production function takes the CES form and let $n(\varphi)$ be the number of suppliers a firm with productivity φ samples. Letting s^i_{jk} be the j'th order statistic of firm i's expenditure share for product k, $\xi^{n(\varphi)}_{ck}$ be the maximum draw of $\left[\xi^f_{ck}\right]^{n_i}_{f=1}$ and $\xi_{jk}(n(\varphi))$ be the j'th order statistic of the distribution of $\left(\xi^{n(\varphi)}_{1k}, \xi^{n(\varphi)}_{2k}\right)$, (18) implies that

$$E\left[ln\left(s^i_{2k}/s^i_{1k}\right)\right] \propto E\left[ln\left(\xi_{2k}(n(\varphi))/\xi_{1k}(n(\varphi))\right)\right]. \tag{25}$$

Our empirical results imply that $E\left[ln\left(\xi_{2k}(n(\varphi))/\xi_{1k}(n(\varphi))\right)\right]$ is decreasing in $n(\varphi)$.[63] To see that property fully depends on the underlying distribution $G(\xi)$, consider the following two examples.

Example 1: Frechet distribution as in Eaton and Kortum (2002) Suppose all suppliers provide the same quality (normalized to unity) of the input but that prices (as in Eaton and Kortum (2002)) are inversely proportional to efficiency z, which itself follows a Frechet distribution, i.e. $F_c(z) = e^{-T_c z^{-\theta}}$. Hence, importing firms will buy from the firm with highest efficiency. As the Frechet distribution is max-stable, a firm who samples n suppliers draws the *selected* level of efficiency the actual supplier in that country from the Frechet distribution $F_c(z|n) = e^{-nT_c z^{-\theta}}$. As price adjusted qualities ξ are proportional to p^{-1}, which in turns are proportional to z, the ratio of expenditure shares (25) is simply determined by the ratio of the highest and second highest efficiency across countries. However, Bernard, Eaton, Jensen, and Kortum (2003) show that this ratio only depends on θ and not on the mean nT_c.[64] Hence: (25) is independent of n and the asymmetric search intensities across firms do not translate into systematic differences in the allocation of spending.

Example 2: Bernoulli distribution Suppose to the contrary that efficiencies are drawn from a Bernoulli distribution where z can be of high (low) quality z_H (z_L) with probability $1-\theta$ (θ). Suppose that $z_L = 1$. The distribution of quality of a supplying country for a firm sampling n suppliers is also Bernoulli and follows the distribution

$$P(z) = \begin{cases} 1-\theta^n & \text{if } z = z_H \\ \theta^n & \text{if } z = z_L = 1 \end{cases}.$$

Again, price adjusted qualities are proportional firm efficiencies so that (25) is given by

$$E\left[ln\left(\xi_{2k}(n(\varphi))/\xi_{1k}(n(\varphi))\right)\right] = -ln(z_H)\, 2\theta^n(1-\theta^n), \tag{26}$$

which reflects that fact that $\theta^n(1-\theta^n)$ is the probability that the importer choses *different* suppliers in both countries and this configuration occurs in two permutations. It is easy to see that (26) is

[63] We for simplicity think search intensity as being governed by the number of suppliers a firm samples. In general, it is more appropriate to think of a stopping rule $\overline{\xi}(\varphi)$, where $\overline{\xi}'(\varphi) > 0$, i.e. bigger firms are more selective. We are only interested to show that - without more restrictions on G - a process of search will not automatically generated non-homothetic demand system.

[64] Note that it is *not* essential that n is the same across countries - the location shifter T_c is country-specific.

decreasing in n (and hence productivity) as long as $n < \frac{ln(1/2)}{ln(\theta)}$. Hence, if high quality suppliers are a rare event ($\theta \approx 1$), more productive firms, which sample more intensely, will have a higher concentration of expenditure and their demand appears to be non-homothetic.

6.8 Controlling for intra-firm trade

Table 12 replicates our analysis for the sample of firm who report having no foreign affiliates abroad. The point estimates are almost identical to our main specifications. In particular, we still find robust evidence for the non-homotheticity of import demand.

	Dependent variable:									
	$ln(s_{ik}^{max})$		$\ln(s_2/s_1)$		$\ln(s_3/s_1)$		$\ln(s_4/s_1)$		$\ln(s_5/s_1)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ln Sales	0.010***	0.010***	-0.101***	-0.104***	-0.149***	-0.155***	-0.171***	-0.163***	-0.178	-0.193
	(0.001)	(0.001)	(0.004)	(0.005)	(0.016)	(0.019)	(0.048)	(0.058)	(0.122)	(0.152)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Fixed Effects	Product-Sourcing Strategy									
N	435512	403438	435512	403438	164901	152751	78107	72378	41782	38754
R^2	0.74	0.74	0.59	0.6	0.83	0.84	0.93	0.93	0.97	0.97

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects, 8 digit product fixed effects and fixed effects for individual sourcing strategies. We weigh observations such that each firm has an equal weight. In columns 1 and 2, the dependent variable is $ln\left(s^i_{max,k}\right)$, where $s^i_{max,k} = \max_{c \in C^i_k} s^i_{ck}$ is the expenditure share on the most popular variety of product k for firm i. In the remaining columns, the dependent variables are $ln\left(s^i_j/s^i_1\right)$, which is the expenditure share on firm i's jth variety relative to its first variety, where the different varieties are ranked by their expenditure shares. All columns with controls control for firms' export status, capital per worker and for indicator variables if the firm is member of a foreign or corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment. We only include firms that do not have any affiliate abroad.

Table 12: Firm Characteristics and the Intensive Margin of Trade: Controlling for Intra-firm Trade

7 Online-Appendix (Not for Publication)

This appendix contains various robustness checks and additional results.

7.1 Robustness of the Main Results to Other Measures of Productivity

In this section we show that our main results are robust with respect to other measures of productivity and other controls for firms' sourcing strategy. In our main analysis, we consistently used sales as our proxy for productivity φ. We think that this choice most closely resembles the logic of the theory. In the following we redo the analysis using estimated productivity as the independent variable and show that all our results are robust. A word of caution is in order. Our data does not contain firm-specific prices. Hence, our productivity measures are purely revenue-based. Without putting more structure on the demand side of the model, there is no reason to believe that measured productivity is related to the model's physical productivity φ (Foster, Haltiwanger, and Syverson, 2008). In fact, in our data, revenue-based measures of productivity are only weakly related to import status unless firm-fixed effects are included in the regression. Hence, we prefer sales as a measure and include the following results for completeness as revenue-based productivity measures have been used in the literature before (e.g. Bernard, Jensen, Redding, and Schott (2007)). We consider two different measures. First we take an OLS-based measure by calculating TFP as the residual from industry-specific OLS regressions of log value-added on log capital and log employment (see Bernard and Jensen (1999); Kugler and Verhoogen (2009)). Second we estimate firm-level TFP using the procedure proposed by Levinsohn and Petrin (2003), and we rely on the specification based on value added, using intermediate inputs as an instrument.

7.1.1 Robustness of Tables 3

Tables 13 and 14 below replicate the results of Table 3 using the different productivity measures as independent variables (Table 13) and using the level of expenditure shares as dependent variable (Table 14). All coefficients are highly significant and have the expected sign.

7.1.2 Robustness of Tables 5 and 24

Tables 15 and 16 replicate the results of Tables 5 and 24, when we measure productivity φ by measured TFP instead of sales. The results are very similar to the ones reported in the main text. All coefficients have the expected sign and are significant for the case of two and three varieties (5) and four varieties (Table 24). For the case of five varieties there are generally too few degrees of freedom to estimate the coefficients precisely.

7.1.3 Robustness of Table 7

Table 17 replicates the results of Table 7, for our two TFP measures as independent variables. All coefficients are of the expected sign and most of them are significant. In Panel A, we control for firms' sourcing strategy by the number of products sourced and find the same asymmetry of the

	Dep. Variable: Max. expenditure share $ln(s_{ik}^{max})$				
	(1)	(2)	(3)	(4)	(5)
	Panel A: TFP (OLS based measure)				
ln TFP	0.003***		0.007***	0.006***	0.008***
	(0.001)		(0.001)	(0.001)	(0.001)
ln Nb. Varieties		-0.280***	-0.280***	-0.283***	-0.287***
		(0.001)	(0.001)	(0.001)	(0.002)
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
Firm Controls	No	No	No	Yes	Yes
N	635,214	658,505	635,214	591,138	591138
R^2	0.06	0.22	0.22	0.23	0.61
	Panel B: TFP (Levinsohn-Petrin based measure)				
ln TFP	0.004***		0.005***	0.006***	0.012***
	(0.000)		(0.000)	(0.000)	-0.001
ln Nb. Varieties		-0.280***	-0.280***	-0.284***	-0.288***
		(0.001)	(0.001)	(0.001)	-0.002
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
Firm Controls	No	No	No	Yes	Yes
N	626,403	658,505	626,403	586,110	586110
R^2	0.06	0.22	0.22	0.23	0.61

	Dep. Variable: Min. expenditure share $ln(s_{ik}^{max})$				
	(6)	(7)	(8)	(9)	(10)
	Panel C: TFP (OLS based measure)				
ln TFP	-0.118***		-0.075***	-0.096***	-0.106***
	(0.005)		(0.004)	(0.004)	-0.009
ln Nb. Varieties		-3.337***	-3.337***	-3.279***	-2.963***
		(0.006)	(0.006)	(0.006)	-0.012
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
Firm Controls	No	No	No	Yes	Yes
N	635,214	655,648	635,214	591,138	591138
R^2	0.09	0.44	0.44	0.45	0.77
	Panel D: TFP (Levinsohn-Petrin based measure)				
ln TFP	-0.073***		-0.053***	-0.062***	-0.142***
	(0.003)		(0.002)	(0.003)	-0.008
ln Nb. Varieties		-3.337***	-3.333***	-3.276***	-2.955***
		(0.006)	(0.006)	(0.006)	-0.012
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
Firm Controls	No	No	No	Yes	Yes
N	626,403	655,648	626,403	586,110	586110
R^2	0.09	0.44	0.44	0.45	0.77

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. We weigh observations such that each firm has an equal weight. In Panels A and B, the dependent variable is $ln\left(s_{ik}^{max}\right)$, where $s_{ik}^{max} = \max_{c \in C_k^i} s_{ck}^i$ is the expenditure share on the most popular variety of product k for firm i. In Panels C and D, the dependent variable is $ln\left(s_{ik}^{min}\right)$, i.e. the expenditure share on the least popular variety. The number of varieties is the number of countries where product k is sourced from. TFP is either the residual of industry specific OLS regressions of log value-added on log capital and log employment (Panels A and C) or estimated using the procedure proposed by Levinsohn and Petrin (2003) (Panels B and D), where we rely on the specification based on value added, using intermediate inputs as an instrument. Columns (5) and (10) include a full set of interacted fixed effects at the product-country-industry level. Columns (4), (5), (9) and (10) control for export status and capital intensity and indicator variables if the firm is member of a foreign or corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment.

Table 13: Robustness of Table 3: Proxy φ by measured TFP

Panel A	Dep. Variable: Max. expenditure share s_{ik}^{max}				
	(1)	(2)	(3)	(4)	(5)
ln Sales	-0.001***		0.008***	0.007***	0.009***
	(0.000)		(0.000)	(0.000)	0
ln Nb. Varieties (prod.)		-0.177***	-0.184***	-0.183***	-0.190***
		(0.001)	(0.001)	(0.001)	-0.001
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
Firm Controls	No	No	No	Yes	Yes
N	655,648	655,648	655,648	594,903	594903
R^2	0.06	0.20	0.20	0.21	0.6

Panel B	Dep. Variable: Min. expenditure share s_{ik}^{min}				
	(6)	(7)	(8)	(9)	(10)
ln Sales	-0.013***		-0.005***	-0.004***	-0.005***
	(0.000)		(0.000)	(0.000)	(0.000)
ln Nb. Varieties (prod.)		-0.169***	-0.165***	-0.163***	-0.157***
		(0.000)	(0.000)	(0.000)	(0.001)
Fixed Effects	Product	Product	Product	Product	Product × Industry × Country
Firm Controls	No	No	No	Yes	Yes
N	655,648	655,648	655,648	594,903	594,903
R^2	0.08	0.26	0.26	0.26	0.69

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. We weigh observations such that each firm has an equal weight. In panel A the dependent variable is s_{ik}^{max}, where $s_{ik}^{max} = \max_{c \in C_k^i} s_{ck}^i$ is the expenditure share on the most popular variety of product k for firm i. In panel B, the dependent variable is s_{ik}^{min}, i.e. the expenditure share on the least popular variety. The number of varieties is the number of countries where product k is sourced from. Columns (5) and (10) include a full set of interacted fixed effects at the product-country-industry level. Columns (4), (5), (9) and (10) control for export status and capital intensity and indicator variables if the firm is member of a foreign or corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment.

Table 14: Robustness of Table 3: Expenditure shares as dependent variables

	$\ln(s_2/s_1)$		$\ln(s_3/s_1)$		$\ln(s_4/s_1)$		$\ln(s_5/s_1)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	LP	OLS	LP	OLS	LP	OLS	LP
	Firms sourcing the exact same 2 varieties of product k							
ln TFP	-0.078***	-0.049***						
	(0.015)	(0.010)						
N	71,477	70,901						
R^2	0.24	0.24						
	Firms sourcing the same exact 3 varieties of product k							
ln TFP	-0.110***	-0.091***	-0.199***	-0.146***				
	(0.033)	(0.022)	(0.044)	(0.030)				
N	20,831	20,614	20,831	20,614				
R^2	0.51	0.51	0.54	0.54				
	Firms sourcing the exact same 4 varieties of product k							
ln TFP	0.029	-0.024	-0.012	-0.055	-0.079	-0.065		
	(0.064)	(0.044)	(0.079)	(0.057)	(0.106)	(0.076)		
N	9676	9564	9676	9564	9676	9564		
R^2	0.76	0.76	0.76	0.76	0.77	0.77		
	Firms sourcing the exact same 5 varieties of product k							
ln TFP	-0.023	-0.06	-0.1	-0.109	0.006	-0.084	-0.078	-0.208
	(0.107)	(0.082)	(0.142)	(0.117)	(0.162)	(0.13)	(0.189)	(0.159)
N	5835	5772	5835	5772	5835	5772	5835	5772
R^2	0.88	0.88	0.88	0.88	0.88	0.88	0.87	0.87

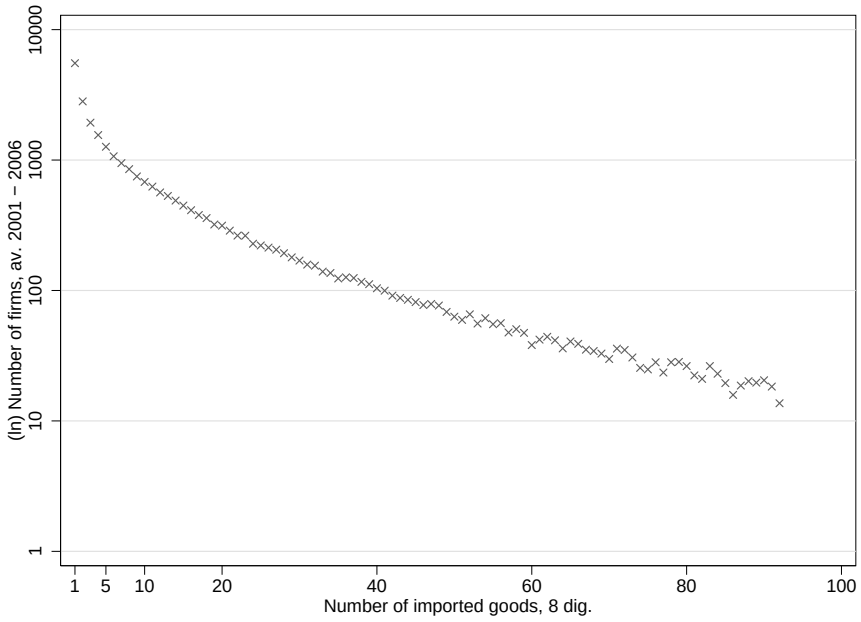
Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects, 8 digit product fixed effects and control for firms' export status and (log) capital intensity and contain indicator variables if the firm is member of a foreign or a corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment. We weigh observations such that each firm has an equal weight. The dependent variables are $ln\left(s_j^i/s_1^i\right)$, which is the expenditure share on firm i's jth variety relative to its first variety, where the different varieties are ranked by their expenditure shares. The table contains 4 sets of regressions, each corresponding to a different number of varieties in the sourcing strategy. In any regression, all firms share the exact same sourcing strategy for a given product (see main text for description of this procedure). In the odd columns, we measure TFP by the residual of industry specific OLS regressions of log value-added on log capital and log employment. In the even columns we estimate TFP using the procedure proposed by Levinsohn and Petrin (2003), where we rely on the specification based on value added, using intermediate inputs as an instrument.

Table 15: Firm Characteristics and Relative Expenditure Shares: Proxying φ by measured TFP

	$\ln(s_2/s_1)$		$\ln(s_3/s_1)$		$\ln(s_4/s_1)$		$\ln(s_5/s_1)$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	LP	OLS	LP	OLS	LP	OLS	LP
Firms sourcing the exact same 2 varieties of product k								
ln TFP	-0.096***	-0.059***						
	(0.010)	(0.006)						
N	170,658	169,130						
R^2	0.15	0.15						
Firms sourcing the same exact 3 varieties of product k								
ln TFP	-0.041***	-0.043***	-0.077***	-0.092***				
	(0.014)	(0.009)	(0.020)	(0.013)				
N	62,517	61,851	62,517	61,851				
R^2	0.27	0.27	0.30	0.30				
Firms sourcing the exact same 4 varieties of product k								
ln TFP	-0.009	-0.022*	-0.030	-0.048***	-0.083***	-0.112***		
	(0.019)	(0.013)	(0.024)	(0.016)	(0.031)	(0.021)		
N	31,198	30,836	31,198	30,836	31,198	30,836		
R^2	0.40	0.40	0.42	0.42	0.42	0.42		
Firms sourcing the exact same 5 varieties of product k								
ln TFP	-0.018	-0.009	-0.027	-0.038*	-0.01	-0.048*	-0.063	-0.093***
	(0.025)	(0.017)	(0.032)	(0.022)	(0.038)	(0.027)	(0.047)	(0.033)
N	18189	17970	18189	17970	18189	17970	18189	17970
R^2	0.54	0.54	0.54	0.54	0.54	0.54	0.53	0.53

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects, 8 digit product fixed effects and control for firms' export status and (log) capital intensity and contain indicator variables if the firm is member of a foreign or a corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment. We weigh observations such that each firm has an equal weight. The dependent variables are $ln\left(s_j^i/s_1^i\right)$, which is the expenditure share on firm i's jth variety relative to its first variety, where the different varieties are ranked by their expenditure shares. The table contains 4 sets of regressions, each corresponding to a different number of varieties in the sourcing strategy. In any regression, all firms e.g. source from the same 2 varieties but might also source it from additional varieties (see main text for description of this procedure). In the odd columns, we measure TFP by the residual of industry specific OLS regressions of log value-added on log capital and log employment. In the even columns we estimate TFP using the procedure proposed by Levinsohn and Petrin (2003), where we rely on the specification based on value added, using intermediate inputs as an instrument.

Table 16: Firm Characteristics and the Intensive Margin of Trade: Exploiting the CES Structure and proxying φ by measured TFP



Notes: The figure shows the number of firms importing k products for different numbers of k. A product is defined on the 8-digit level and a firm is defined to be an importer of product k, whenever it imports a positive amount from at least one country. We use 6 years of data from 2001-2006 and report the yearly average.

Figure 5: How many products do firms import?

intensive and extensive margin as reported above. These result are highly significant. In Panel B we perfectly control for firms' sourcing strategy. As in Table 7, some results are not significant because we have too few degrees of freedom.

7.2 Variety-Facts at the Product Level

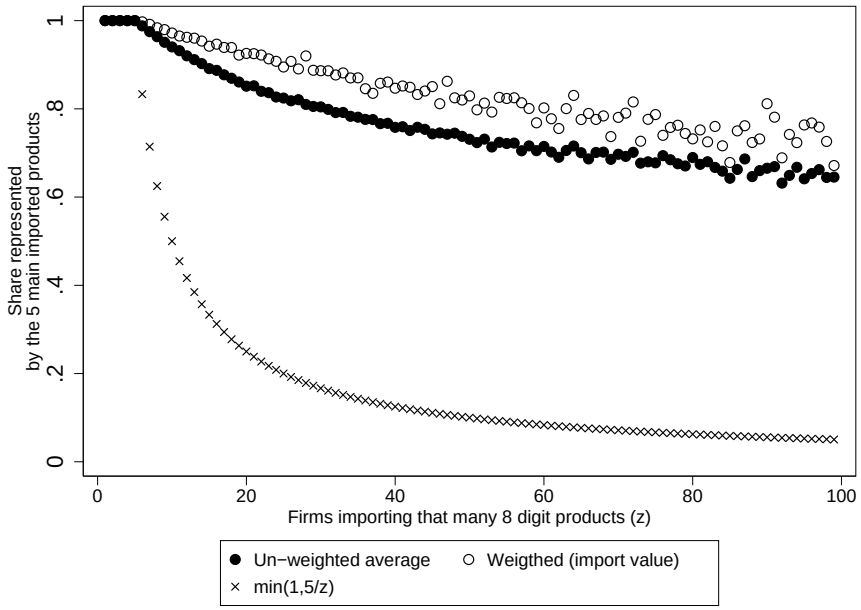
Here we show that the two facts documented in Section 3.1 at the variety level also hold at the product level. Figure 5 shows the number of firms importing k products for different values of k. As for the case of varieties, there is substantial heterogeneity. Whereas roughly 20% of firms source only one product, there are a number of firms importing more than 40 products from abroad. On average, French importers import 13 products from abroad, but due to the skewness of the distribution the median number of products is only 6. Interestingly, the graph is almost linear, so that the semi-elasticity of the "product extensive margin" is roughly constant. Increasing the number of imported products by one decreases the number of firms doing so by 1.4 percent. Figure 6 turns to the intensive margin at the product level. As for the case of varieties, expenditure share are very concentrated: firms sourcing 50 product still allocate 80% of their expenses on merely 5 products and even firms

	Dep. variable: log Average price ($\omega_k^i = \sum_{c=1}^{C_{k,i}} p_{ck}^i s_{ck}^i$)							
Panel A	Firms sourcing at least 2 varieties of product k							
	OLS based measure				Levinsohn-Petrin based measure			
ln TFP	0.051***		0.052***	0.061***	0.028***		0.028***	0.030***
	(0.003)		(0.003)	(0.003)	(0.002)		(0.002)	(0.002)
ln nb. varieties (prod.)		-0.021***	-0.024***	-0.039***		-0.021***	-0.023***	-0.039***
		(0.004)	(0.004)	(0.004)		(0.004)	(0.004)	(0.004)
Firm Controls	No	No	No	Yes	No	No	No	Yes
N	591,340	610,375	591,340	549,262	582,909	610,375	582,909	544,432
R^2	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74

Panel B	Firms sourcing the exact same ... of product k							
	2 varieties		3 varieties		4 varieties		5 varieties	
	OLS	LP	OLS	LP	OLS	LP	OLS	LP
ln TFP	0.035***	0.021***	0.027	0.021	0.066	0.068**	0.002	0.001
	(0.008)	(0.006)	(0.018)	(0.014)	(0.045)	(0.033)	(0.074)	(0.065)
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	75,923	68,309	22,287	20,072	10,426	9,356	6,270	5,654
R^2	0.86	0.86	0.94	0.94	0.97	0.97	0.99	0.99

Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. All regressions include year fixed effects, 3 digit industry fixed effects and 8 digit product fixed effects. Specification with firm controls control for (log) capital intensity and export status and contain indicator variables if the firm is member of a foreign or a corporate group. A firm is member of a foreign group if at least one affiliate or the headquarter is located outside of France. A firm is member of a corporate group if it is controlled by another firm or it has control over at least one affiliate. Capital per worker is measured as tangible capital (in thousand euros), as reported in the firm's balance sheet, divided by employment. In Panel A we consider all firm-product pairs, who source at least two varieties of the respective product. In Panel B we consider all firm-product pairs, who source a particular from the exact same sourcing countries. We measure TFP either by the residual of industry specific OLS regressions of log value-added on log capital and log employment or by the procedure proposed by Levinsohn and Petrin (2003), where we rely on the specification based on value added, using intermediate inputs as an instrument.

Table 17: Firm Size and Average Import Prices: Proxying φ by measured TFP



Notes: The figure shows the average expenditure share on the 5 most popular products of firms importing k products for different values of k. For the unweighted average all firms get an equal share, for the weighted results, we weigh firms by their import value. We also depict the counterfactual expenditure share if expenditures were equalized across products.

Figure 6: Concentration of firms' import spending across products.

that source 100 products do so with 70% of their budget.