Article:
Navas, A., Serti, F. and Tomasi, C. (2020) The role of the gravity forces on firms' trade. Canadian Journal of Economics, 53 (3). pp. 1059-1097. ISSN 0008-4085

https://doi.org/10.1111/caje.12467

The role of the gravity forces on firms' trade

Antonio Navas[a], Francesco Serti[b], Chiara Tomasi[c,*]

[a]*Department of Economics, University of Sheffield*
[b]*FAE, Universidad de Alicante and AXES IMT, Lucca*
[c]*Department of Economics and Management, University of Trento and LEM, Pisa*

Abstract

This paper offers both a theoretical framework and empirical evidence on the role that the two gravity forces, namely market size and geographical distance, have indirectly through imports, on firms' exports patterns. The model shows that sourcing from bigger and closer markets implies higher productivity gains which, in turn, increase firms' ability to enter export market, as well as their export value. Exploiting data on product and destination-level transactions of a large panel of Italian firms, the paper shows that, on average, the indirect effects of the gravity forces are about one third of their direct effects.

JEL codes: F12, F14.
Keywords: Gravity Forces, Geographical distance, Market size, Firms' heterogeneity, Imports and Exports.

---

*Corresponding author. *Email:* chiara.tomasi@unitn.it, *Address*: Department of Economics and Management, Via Inama 5, 38122, Trento, Italy. *Tel* +39 (0)461 282161

Email addresses: a.navas@sheffield.ac.uk (Antonio Navas), serti@gmail.com (Francesco Serti)

1. Introduction

The gravity equation in international trade has been largely used as a workhorse for analyzing the determinants of bilateral trade and the geographical patterns of economic activity. The equation shows that trade between two countries is proportional to their respective sizes and inversely proportional to the geographic distance between them. The recent heterogeneous-firms trade models have shown that the gravity forces shape export performance also at the firm-level: larger size or lower distance increase the probability that a firm exports to a particular destination (the extensive margin), as well as its export value to that market (the intensive margin) (Chaney, 2008; Helpman et al., 2008).[1] The same literature has shown that there exist a strong complementarity between firms' import and export activities: the majority of exporters are also importers and vice versa. The connection between the two sides of trade could have important consequences for the elasticity of exports with respect to the gravity forces (Bernard et al., 2016).

This paper improves upon the existing literature by considering the role that the gravity forces have on firms' export patterns through the imports of intermediate inputs. Our research provides empirical evidence that market size and geographical distance have an indirect effect, through imports, on a firm' s probability of exporting and its export value. As in previous papers, in our setting sourcing intermediate inputs from abroad has a positive impact on firms' exports because of productivity enhancing effect due to variety, quality and technological mechanisms.[2] This work contributes to the existing literature by showing that the economic geography of imports is crucial in influencing firms' productivity and exports: productivity gains from imports depend indeed on the size and the distance of the source countries. As a consequence, the elasticity of exports with respect to gravity forces is magnified.

To guide our empirical analysis we introduce in the theoretical framework of Chaney (2008), which derives the export gravity equation for final goods in a model of trade with firm heterogeneity, an intermediate input sector and the possibility for final producers to use a continuum of intermediate inputs sourced from multiple locations differing in terms of size, labour costs, trade and institutional barriers. The technology exhibits love of variety in intermediate inputs in the spirit of early endogenous growth models (Romer, 1990; Rivera-Batiz and Romer, 1991). The model predicts that the positive effect of importing on a firm's productivity is heterogeneous across source countries and it depends on both the mass of imported intermediate inputs available, as well as the price of each intermediate. Bigger markets provide a larger variety of inputs, while closer countries charge lower prices because of lower transportation costs. Therefore, variation in the gravity forces determine heterogeneous productivity gains across import-source countries: importing from larger and closer markets has a stronger positive effect on firms' productivity. The efficiency gains, in turn, increase a firm's probability of exporting, as well as its export value. It follows that, in addition to the direct effect on firms' export patterns, market size and distance exert an effect on exports indirectly through heterogeneous efficiency gains induced by imports of intermediate inputs. A decline in transportation costs (i.e., distance), and therefore a reduction in the cost of imported inputs, increases a firm's productivity allowing it to offer its exports at lower prices and to increase its revenues in the

[1]Head and Mayer (2014) propose a review of the estimation and interpretation of the gravity equation for bilateral trade.

[2]For a theoretical background of the productivity gains induced by intermediate inputs see Markusen (1989); Grossman and Helpman (1991); Acharya and Keller (2009); Eaton et al. (2011), among others. Micro-level empirical evidence on the positive effect of imports on firms' productivity include Halpern et al. (2015) for Hungary, Paul and Yasar (2009) for Turkey, Conti et al. (2014) for Italy, Gorg et al. (2008) for Ireland, Vogel and Wagner (2010) for Germany. Other relevant papers that investigate the effect of input trade liberalization on firms' productivity are Fernandes (2007); Pavcnik (2002); Amiti and Konings (2007); Kasahara and Rodrigue (2008), among others. Alternatively, the link between importing and exporting could be due to the existence of sunk costs complementarities between the two activities as in Kasahara and Lapham (2013) (see section 4.2.1 for a discussion on this channel).

exporting markets.[3] Following a similar reasoning, the bigger the foreign country, the larger the mass of imported inputs and the lower the marginal cost of production: a rise in the size of the foreign market determines larger efficiency gains and thereby increases a firm's export performance.

The theoretical set-up helps us in driving our empirical analysis. We exploit an original Italian database obtained by merging a firm-level dataset, including standard balance sheet information, with a transaction-level dataset, recording custom information on exports and imports for each product and destination. Firm-level trade data are complemented by country characteristics including proxies for market size, distance, variable and fixed trade costs. We adapt the gross output production function estimation method proposed by Gandhi et al. (2018) to our theoretical setting by taking into account the role of imports of intermediate inputs and we derive the contribution of imports to a firm's total factor productivity. Our results point at the importance of foreign intermediates in explaining productivity differences across firms within sectors. On average our estimates indicate that a firm that increases its ratio of total intermediate inputs (foreign plus domestic) over domestic intermediate inputs by 10% can improve its total factor productivity (TFP) by 3.2%.

We then test for the indirect effect that the two gravity forces, through the import-related component of TFP, have on a firm's export participation and export sales in a destination market. We adopt an instrumental variable (IV) strategy to control for possible endogeneity bias of our key variable due to omitted variables or reverse causality. The empirical analysis provides evidence that the firms' productivity component due to imports, which is heterogeneous across import-source countries, has a positive impact on both firm-country export margins. We also confirm previous empirical results (Lawless and Whelan, 2014; Crozet and Koenig, 2010) regarding the direct channel of the two gravity forces according to which the probability of exporting to a specific market, as well as the amount of exports, increase with market size but decrease with distance. Finally, we quantify the indirect effects of the two gravity forces on a firm's export patterns. We estimate that on average the indirect effects of the gravity forces are about one third of the direct effects obtained in the gravity equations.

Our paper directly relates to the literature on the gravity equation. Applied for the first time by Tinbergen (1962), the equation shows that trade between two countries is proportional to their respective sizes, measured by their GDP, and inversely proportional to the geographic distance between them. The heterogeneous-firm model brings to the gravity model a need to consider the effects of trade barriers both on the value of exports by current exporters and on the entry of exporters. In his model Chaney (2008) extends the work of Melitz (2003) to show that there is both an intensive and an extensive margin of adjustment of trade flows to trade barriers. In a similar manner, Helpman et al. (2008) derive a gravity equation and develop an estimation procedure to obtain the effects of trade barriers and policies on the two margins. Micro-level empirical analyses confirm several of the theoretical implications predicted by these models. Eaton et al. (2011, 2004) for France and Bernard et al. (2007) for the US find that the number of exporting firms is sharply decreasing in the distance to the destination country and increasing in importers' income. Using French firm-level data, Crozet and Koenig (2010) estimate the effect of trade barriers on different export margins. Other empirical studies offer evidence that market-specific trade costs affect individual export decision and export sales to a particular destination (Lawless and Whelan, 2014; Creusen et al., 2011; Serti and Tomasi, 2014). By considering the import side, Loof and Andersson (2010) and Conti et al. (2014) estimate the causal impact of importing from different sources on a firm's productivity.

Within the vast theoretical and empirical literature on firm heterogeneity in international trade, this article directly relates to the emerging literature on the interdependence between im-

[3]The result that intermediates magnify the elasticity of trade flows to trade barriers is also provided by Yi (2003); Caliendo and Parro (2015); Aichele et al. (2014). Our theoretical framework emphasizes the role of firm heterogeneity and of self-selection across both export and import activities.

porting and exporting activities. From a theoretical point of view Kasahara and Lapham (2013) builds a symmetric two-country model on the import-productivity-export nexus. Halpern et al. (2015) includes multiple intermediate goods and a fixed cost of importing which is increasing in the number of intermediate inputs imported. Their paper mainly focuses on estimating the impact of importing on TFP and determining the share of the productivity gains that comes from an increase in the quality of intermediate inputs and the share that comes from an imperfect substitution between foreign and domestic inputs. The positive link between imports and exports is confirmed empirically by Bernard et al. (2007) for the US, Bas and Strauss-Kahn (2014) for France, Feng et al. (2016) for China, Muuls and Pisu (2009) for Belgium, LoTurco and Maggioni (2015) for Turkey, Aristei et al. (2013) for a group of Eastern European and Central Asian countries, Altomonte and Bekes (2009) for Hungary, Kasahara and Lapham (2013) for Chile. Evidence for Italy has been provided by Castellani et al. (2010) and LoTurco and Maggioni (2013). Other papers look at the connection between the two trade activities by investigating the effect of input-trade liberalization on firm export outcomes (Bas, 2012; Bas and Strauss-Kahn, 2015; Fan et al., 2016; Chevassus-Lozza et al., 2013).

Our work departs from those cited above and it contributes to the existing literature by considering how the economic geography of import activities impacts exports trough its heterogeneous effect on firms' productivity. It therefore differentiates with respect to previous works (e.g. Kasahara and Lapham, 2013; Halpern et al., 2015) by considering not only how imports affect a firm's productivity but also how the gravity forces influence this import-productivity nexus and, in turn, the import-export linkages. While it has been already established that market size and distance are crucial in shaping exports patterns, it is an open question whether and how the two gravity variables play a role indirectly through imports. In order to explore the role played by the gravity variables on the gains from importing and the import-export productivity nexus, a dimension so far unexplored in the literature, the theoretical model adapts the existing frameworks to a multi-country setting. Unlike the previous settings, our model suggests that the gains from importing are expected to be heterogeneous across source countries with higher gains coming from closer and larger source countries. The empirical part corroborates the predictions of the theory and it provides unique estimates for each source country on the impact of importing intermediates on a firm's export patterns.

The remainder of the paper is organized as follows. Section 2 presents a trade model with heterogeneous firms, featuring imports in intermediate inputs to derive the export gravity equation, both at firm and industry-level. Section 3 describes the data for the empirical study. Section 4 presents the estimation results and Section 5 concludes.

2. The model

We develop a simple model that extend Melitz (2003) and Chaney (2008) to incorporate trade in intermediates in an asymmetric countries environment. Our aim is to derive the firm-country equations for export participation and values which include cross country determinants of export and import activities, which is the focus of the paper.

2.1. The firm-level export gravity equation

We consider a model with N potential asymmetric countries, indexed by n, each of them populated by a continuum of individuals of measure L_n. Individuals derive utility from the consumption of the $H+1$ final goods existing in the economy, with Q_{hn} representing consumption of final good h in the generic country n and μ_h is the optimal share of expenditure devoted to good h. Sector 0 produces an homogeneous good while each of the rest H different sectors produces a continuum of varieties ω in the set Ω_h. Preferences across different varieties of the same final good are described by standard CES utility function with $q_{hn}(\omega)$ denoting the quantity consumed of variety ω of good h in country n and the parameter σ_h controls for the elasticity of substitution across varieties within the sector h.

The homogeneous good is produced under perfect competition using a linear technology. To produce one unit of the homogeneous good in country n, a firm needs to employ $(1/\iota_n)$ units of labour. As standard in this literature, if we consider this good as the numeraire, perfect competition implies that this sector pins down the wages in each country (i.e. $w_n = \iota_n$).

In the other final good sectors, each firm produces a unique differentiated variety. To produce, each firm f in sector h needs to incur in a per period fixed cost of operation F_h (in units of the numeraire). We assume that firms use intermediate inputs and labor to produce using the following Cobb-Douglas technology

$$q_{hn}^f = \varphi_h^f \left(l_{hn}^f\right)^{1-\alpha_h} \left(m_{hn}^f\right)^{\alpha_h} \tag{1}$$

where l_{hn}^f denotes labor dedicated to production, $m_{hn}^f = \left(\int\limits_{\nu\epsilon\Lambda} \left(m_{hn}^f(\nu)\right)^{\frac{\phi_{h-1}}{\phi_h}} d\nu\right)^{\frac{\phi_h}{\phi_h-1}}$ is the intermediate composite input used in sector h where $m_{hn}^f(\nu)$ is a firm f's demand of the intermediate input variety ν produced in country n, and φ_h^f denotes a firm's innate productivity described below. The parameter $\phi_h > 1$ controls for the degree of substitutability across intermediate inputs within a sector. The parameter α_h measures the importance of intermediate inputs in the production of each final good. Both are assumed to be identical across countries. Common to Romer (1990) and Rivera-Batiz and Romer (1991), we assume that there are diminishing marginal returns associated with each intermediate input variety (i.e., $\frac{\phi_h-1}{\phi_h} < 1$) while a firms' production increases with the mass of varieties of intermediates used.

Firms in the H final good sectors differ in their innate productivity φ_h^f. Following Chaney (2008), we assume that each firm, at the moment of entry, obtains its innate ability from a common Pareto distribution with cumulative distribution function given by

$$\Pr(\varphi_h^f \leq \varphi) = 1 - \varphi^{-\gamma_h}$$

with γ_h controlling for the productivity dispersion within sectors.[4]

In the intermediate input sector, each firm within each country is producing a unique variety using one unit of labor to produce one unit of output. As in Chaney (2008), we assume that the mass of entrants is proportional to $w_n L_n$ (i.e. the labor income of the economy) and we denote with $0 < \beta_{sn} < 1, s = h, m$, respectively the proportion of firms in each final good sector h and in the intermediate input sector m in country n.

Firms can trade in both final goods and intermediate inputs. Moreover, both activities bear fixed and variable costs. A firm in the final good sector h and country k which wants to export (x) to country j must pay a fixed cost of F_{hxkj} units of the homogeneous good, while a firm in the same sector and country k which wants to import (i) needs to pay a fixed cost of F_{hik} units. In order to keep tractability in the model, we assume that once a firm pays F_{hik}, it has access to all the intermediate inputs varieties available in the world.[5] In section 2.2, we show that the effects of importing intermediates on exporting at the firm-level can be summarized with one statistic independently of the latter assumption. The inclusion of fixed costs in both activities implies that not all firms are going to find it profitable either to export final goods or to import intermediates. Therefore, the model predicts self-selection in both exporting and importing activities based on productivity levels. In addition both type of exporters, final good and intermediate producers, bear variable trade costs of the iceberg type (τ). We follow Anderson and van Wincoop (2004) in assuming that the amount of units an exporter must ship for selling one unit of its product to a destination is a log-linear function of the distance

[4]Following the broad literature on trade and firm heterogeneity we assume $\gamma_h > \sigma_h - 1$ and $\gamma_h > 2$.

[5]Antras et al. (2017) provide a multi-country model of importing featuring origin-specific fixed costs of sourcing.

between countries (D) (with an elasticity of trade costs to distance given by δ) and of other variable costs Δ (e.g. tariffs). More precisely,

$$\tau_{xkj} = \Delta_{xkj} \left(D_{kj}\right)^{\delta_x} \tag{2}$$
$$\tau_{mjk} = \Delta_{mjk} \left(D_{kj}\right)^{\delta_m}$$

where τ_{xkj} are the variable trade costs required to export final goods from country k to country j and τ_{mjk} are the variable trade costs required to export intermediate inputs from j to k.

In this model entry is exogenous and firms earn positive profits. To complete the definition of the model, as it is common in the literature, we assume that all existing firms in the world belong to a mutual fund and each individual in each country owns w_n shares of this mutual fund.

Given the general set-up, we can now derive the two firm-level export gravity equations, for the extensive and the intensive margin respectively, which are the focus of the current work. In order to obtain these two expressions one needs to derive the firms' productivity threshold required to survive in the market (φ^*_{hk}), to export to a country j (φ^*_{hxkj}), and to import (φ^*_{hik}). Indeed, the export productivity point, φ^*_{hxkj}, depends on the aggregate price index which is an endogenous variable that, in turn, depends on both the import and the survival productivity thresholds. Using these productivity cutoffs and solving for the aggregate price index allows to obtain the export gravity equations. The details of the derivation for the firm-level extensive and the intensive margin of exports are provided in Section 2.1 of the Technical Appendix.

Since the model is deterministic, depending on the parameters configuration we can have different types of equilibria. Here our focus is on equilibria where the firms engaged in international trade are either both exporters of final goods and importers of intermediate products or just only importers. The sufficient and necessary condition for the existence of this type of equilibria is reported in Section 2.1 of the Technical Appendix.

The firm-level gravity equation for the extensive margin of exports, that is the probability that a firm in country k exports to country j, is given by

$$\Pr(\varphi \geq \varphi^*_{hxkj}) = (\lambda'_{4h})^{-\gamma_h} \underbrace{\left(\frac{Y_j}{Y}\right) \left(\frac{w_k \tau_{xkj}}{\theta'_{hj}}\right)^{-\gamma_h} (F_{hxkj})^{\frac{-\gamma_h}{\sigma_h - 1}}}_{Chaney's} \underbrace{(\tilde{\chi}_{hk})^{\gamma_h}}_{intermediate\ contribution} \tag{3}$$

where λ'_{4h} is a constant, θ'_{hj} is the multilateral resistance term and $\tilde{\chi}_{hk} = \chi_{hk} \left(\frac{\beta_{mk} Y_k}{Y}\right)^{\frac{\alpha_h}{\phi_h - 1}} \left(\frac{1+\pi}{Y}\right)^{\frac{\alpha_h}{1-\phi_h}}$.

This expression relates the standard elements found in a gravity equation to the probability that a firm in k exports to country j (and therefore the mass of firms in k exporting to country j). The last element of equation (3) captures the contribution of intermediate inputs to a firm's Total Factor Productivity (TFP). This element is crucial to our analysis as we will show in the next section.

The firm-level gravity equation for the intensive margin of exports, that is the firm's export volume to country j, is given by

$$X^f_{hxkj}(\varphi^f) = (\lambda'_{3h}) \underbrace{\left(\frac{Y_j}{Y}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \left(\frac{\theta'_{hj}}{w_k \tau_{xkj}}\right)^{\sigma_h - 1}}_{Chaney's} \underbrace{(\tilde{\chi}_{hk})^{\sigma_h - 1}}_{intermediate\ contribution} \left(\varphi^f\right)^{\sigma_h - 1} \tag{4}$$

where λ'_{3h} is a constant.[6] As for the extensive margin, a firm's exports to country j depend on intermediate inputs, as expressed by the intermediate contribution term of equation 4.

[6]Following Chaney (2008)'s notation the variable $\lambda'_{3h} = \sigma_h \left(\lambda'_{4h}\right)^{1-\sigma_h}$ and the variable $\lambda'_{4h} = \left(\frac{\gamma_h}{\gamma_h - (\sigma_h - 1)}\right)^{\frac{1}{\gamma_h}} \left(\frac{\sigma_h}{\mu_h}\right)^{\frac{1}{\gamma_h}} (1+\pi)^{\frac{-1}{\gamma_h}} \psi_h \left(\frac{1+\pi}{Y}\right)^{\frac{\alpha_h}{1-\phi_h}}$.

2.2. Imports, total factor productivity and country characteristics

Given expressions 3 and 4, we derive a set of predictions that can be tested empirically.

Proposition 1. *Importing intermediate inputs has a positive effect on a firm's productivity. This effect depends on the characteristics of the country of origin of imports.*

Since a firms' technology presents diminishing marginal returns associated with each intermediate input variety, importing intermediates allows a firm to escape from these diminishing marginal returns by splitting its intermediate input requirements across more varieties. The ability of a firm to do so depends on the mass of imported intermediate inputs available (i.e., the more varieties available, the better a firm can spread these intermediate input requirements), which is increasing in the size of each source country, as well as on the price of each imported intermediate input, which is decreasing in variable trade costs. Indeed, it is possible to derive a firm's TFP

$$\frac{q^f_{hk}}{\left(l^f_{hk}\right)^{1-\alpha_h}\left(M^f_{tot}\right)^{\alpha_h}} = \varphi^f_h \underbrace{\underbrace{\left[\sum_{j=1}^{N}\left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h}\frac{\beta_{mj}}{\beta_{mk}}\frac{Y_j}{Y_k}\right]^{\frac{\alpha_h}{\phi_h-1}}}_{\chi_{hk}}\left(\frac{\beta_{mk}Y_k}{Y}\right)^{\frac{\alpha_h}{\phi_h-1}}\left(\frac{(1+\pi)}{Y}\right)^{\frac{\alpha_h}{1-\phi_h}}}_{\tilde{\chi}_{hk}} \tag{5}$$

where M^f_{tot} is the total value of (domestic and foreign) intermediate inputs used by a firm deflated by the price of the domestic intermediate input. The left-hand side is the expression for the TFP of a firm f belonging to sector h, that we will bring directly to the data in section 4.1. In the right-hand side of the equation, the first term represents a firm's innate productivity and the second term, $\tilde{\chi}_{hk}$, captures the contribution of intermediate inputs to a firm's TFP. The term $\tilde{\chi}_{hk}$ is a weighted sum of the varieties sourced from each country, where the weights take into account the fact that varieties coming from different countries have different prices.[7] This term can be conveniently decomposed in an element, χ_{hk}, that reflects the gains from importing intermediates, and the other component that account for the gains from variety stemming from domestic intermediate inputs. Note that the variable $\chi_{hk} = 1$ when a firm does not import (and therefore there are no gains associated with importing) and it is $\chi_{hk} > 1$ when a firm imports. As this term enters in a multiplicative way in the expression for the TFP, importing intermediates increases a firm's TFP.

Equation 5 reveals that the gains from importing are decreasing in variable trade costs, τ_{mjk}, and increasing in the economic size of each source country, Y_j. First, trade costs affect the prices of intermediate inputs and, with that, the ability of a firm to reallocate the intermediate input requirements from domestic to foreign varieties. Indeed, when a firm decides the optimal bundle of intermediate inputs to be used, it decides to employ each variety up to the point when the value of its marginal productivity is equal to its price. The existence of diminishing marginal returns associated with each intermediate input variety implies that, when a firm faces the opportunity to import intermediates, it reallocates some of its intermediate inputs requirements from domestic varieties to imported varieties, whose marginal productivity is higher (since they have not been used before). This allows the company to increase its efficiency. The extent to which a firm is able to do this depends on the price of imported intermediates: the cheaper are

[7] We can rewrite this second term, $\tilde{\chi}_{hk}$, as $\left[\sum_{j=1}^{N}\left(\frac{w_j\tau_{mjk}}{w_k}\right)^{1-\phi_h}\beta_{mj}w_jL_j\right]^{\frac{\alpha_h}{\phi_h-1}}$, where $\beta_{mj}w_jL_j$ represents the mass of varieties of intermediate inputs available in country j and the first element in parenthesis indicates the relative price of these varieties with respect to the domestic ones. Note that, if there were no transportation costs and wages were equal across countries this expression will be reduced to $\left[\sum_{j=1}^{N}\beta_{mj}wL_j\right]^{\frac{\alpha_h}{\phi_h-1}}$ which is a sum of all varieties of intermediate inputs available to a firm from its multiple source countries.

the imported intermediates, the more a firm will substitute domestic varieties with foreign ones. Second, the economic size of each country determines the mass of varieties of intermediate inputs offered in the market. Larger countries offer more varieties and this allows a firm to split the intermediate input requirements across more varieties, contributing to larger efficiency gains. Therefore, as commented above, the gains from importing depend *(i)* on the transportation costs and *(ii)* on the economic size. These properties are important for the results found in propositions 2 and 3 below.

An important property of this framework is that the term χ_{hk}, capturing the effect of importing intermediates on a firm's TFP, can be shown to be equal to $\left(\frac{M^f_{tot}}{M^f_k}\right)^{\frac{\alpha_h}{\phi_h - 1}}$, where M^f_k is the total volume of domestic intermediate inputs.[8] This equivalence will allow us to obtain an estimation of the gains from importing, χ_{hk}, from the data.

Note that the result concerning the gains from importing is robust to an alternative richer environment in which a firm bears fixed costs of importing per market, which are source-country specific. When the fixed costs of importing are heterogeneous across countries, a firm's choice regarding the number of source markets will depend on the characteristics of these markets and on its innate productivity. This will influence the number of countries included in χ_{hk}. However, the statistic $\left(\frac{M^f_{tot}}{M^f_k}\right)^{\frac{\alpha_h}{\phi_h - 1}}$ would still capture the positive contribution of importing on a firm's TFP. Therefore, Proposition 1 of the model holds both in the simplified setting of a unique fixed cost of importing and in the more general case in which there are multiple fixed costs of importing and these are heterogeneous across countries.

Proposition 2. *The effect of distance on a firm's probability of exporting and its export value is magnified by the presence of trade in intermediate inputs.*

To the extent that export and import variable costs have common determinants, as assumed in the model, a decrease in transportation costs has a comparatively larger impact on exports than in the absence of intermediate imports. This is the consequence of the fact that a reduction in distance affects a firm's export patterns through a direct effect, standard in the literature, and an indirect effect, via importing. Taking logs and derivatives in equation (3) we obtain the effect that a decrease in D_{kj} has on a firm's export status

$$\frac{d \ln(\Pr(\varphi \geq \varphi^*_{xkj}))}{d \ln(D_{kj})} = -\delta_x \gamma_h + \gamma_h \frac{d \ln \chi_{hk}}{d \ln D_{kj}} \tag{6}$$

and a similar expression is obtained for a firm's export value

$$\frac{d \ln X^f_{hxkj}(\varphi^f)}{d \ln D_{kj}} = -\delta_x (\sigma_h - 1) + (\sigma_h - 1) \frac{d \ln \chi_{hk}}{d \ln D_{kj}}. \tag{7}$$

The direct effect corresponds to the first element on the right hand side of equations (6) and (7). That is, a reduction in the transportation costs between the country of origin k and the country of destination j allows a firm to charge lower prices, increasing both the probability that a firm becomes an exporter to that destination and its export sales to that country. The indirect effect is inherent to this framework and it is captured by the second element of both equations. The reduction in transportation costs between k and j decreases the cost of importing intermediates from country j. This allows a firm to better reallocate its intermediate input requirements across existing varieties and, as a consequence, to become more efficient, as indicated in equation (5). The increase in a firm's TFP allows to charge lower prices, increasing

[8] See section 2.2 in the Technical Appendix for a formal proof.

its probability of exporting and its export sales by the same amount to all destinations (not only to country j).[9]

Proposition 3. *The effect of market size on a firm's probability of exporting and on its export value is magnified by the presence of trade in intermediate inputs.*

Taking logs and derivatives in equation (3) we obtain the effect that a decrease in Y_j has on a firm's export status

$$\frac{d\ln(\Pr(\varphi \geq \varphi^*_{hxkj}))}{d\ln Y_j} = 1 + \gamma_h \frac{d\ln \chi_{hk}}{d\ln Y_j}. \tag{8}$$

and a similar expression is obtained for a firm's export value

$$\frac{d\ln X^f_{hxkj}(\varphi^f)}{d\ln Y_j} = \left(\frac{\sigma_h - 1}{\gamma_h}\right) + (\sigma_h - 1)\frac{d\ln \chi_{hk}}{d\ln Y_j}. \tag{9}$$

An increase in foreign market size has a positive effect on exports due to both a direct and an indirect effect. The first effect, present in Chaney (2008) and Crozet and Koenig (2010) among others, comes from a demand mechanism. *Ceteris paribus*, the larger the economic size of country j, the larger the demand for final goods and therefore the larger the potential sales of exporters. This reduces the productivity level necessary to cover the fixed costs of exporting to that destination and it increases a firm's export sales to that country.[10]

In addition, and novel to this framework, this effect is magnified by the fact that the foreign market is also a source of intermediate inputs. The larger the source country, the larger the mass of imported intermediate inputs. The access to a larger set of intermediate input varieties coming from that country has a positive effect on a firm's TFP, as discussed above, and it allows a firm to charge lower prices. As a consequence, this leads to an increase in a firm's probability of becoming an exporter and in its export value to country j, as well as to all other destinations ($s \neq j$).[11]

The simple theoretical model presented in the paper explores the potential effect of changes in trade costs or market sizes on a firm's export patterns in a tractable manner. The main predictions of the model holds in a more complex but richer environment in which we allow for technological differences in the production of intermediates across countries and differences in the quality of intermediate inputs across source countries. Section 2.3 of the Technical Appendix discusses the robustness of our results under these alternative assumptions.

---

[9]Indeed, the decrease in the cost of importing from country j has the same impact also on a firm's export behavior to destination country s, with $s \neq j$, that is $\frac{d\ln(\Pr(\varphi \geq \varphi^*_{hxks}))}{d\ln(D_{kj})} = \gamma_h \frac{d\ln \chi_{hk}}{d\ln D_{kj}}$ and $\frac{d\ln X^f_{hxks}(\varphi^f)}{d\ln D_{kj}} = (\sigma_h - 1)\frac{d\ln \chi_{hk}}{d\ln D_{kj}}$. An analogous result holds for the impact of foreign market size Y_j.

[10]This result is not present in the Melitz (2003)'s setup with symmetric countries where an increase in market size neither increase a firm's export sales nor the export productivity threshold. In the Melitz (2003) model the potential positive effect of market size on a firm's sales is compensated by the fact that a bigger market size will encourage entry, increasing the mass of varieties available to consumers and reducing each consumer's expenditure in each variety and therefore sales. Thus, the final effect will be the combination of both the positive and the negative effect. The net effect leaves a firm's export sales and the export survival productivity threshold unchanged and it leads to an increase in the number of firms/varieties in the market (as in Krugman (1980)). In contrast, our model, as in Chaney (2008) or Crozet and Koenig (2010), assumes exogenous entry. In the empirical part, we find a direct positive impact of economic size (GDP) on exports at the firm-level at both margins, the extensive and the intensive. A similar empirical result is obtained by Lawless and Whelan (2014).

[11]In this framework the domestic market size also affects a firm's export behaviour. More populated and more productive economies provide a greater number of varieties of intermediate inputs which increases a firm's TFP (this is reflected in equation 5). The increase in a firm's TFP decreases the marginal cost of production which allows a firm to charge lower prices. The latter gives a competitive advantage to domestic firms in foreign markets. Unfortunately, we are not able to test this prediction since we have information only for one domestic market, that is Italy.

3. Data

This section describes the firm-level data and the country-level variables employed in the regressions. The empirical analysis combines two sources of data collected by the Italian Statistical Office (ISTAT): the Italian Foreign Trade Statistics (COE), and a firm-level accounting dataset (Micro.3).[12] The data are available for the period 2000-2006.

The COE dataset is the official source for the trade flows of Italy and it reports all cross-border transactions performed by Italian firms.[13] For all trade flows, we observe annual values, expressed in euros, disaggregated by countries of destination for exports and markets of origin for imports. The available information on product categories, classified according to the 6-digit Harmonized System allows us to single out firms' imports in intermediate inputs defined as those falling into the intermediate input category according to the Broad Economic Categories (BEC) classication of HS6 products. The BEC classification has been widely used in the literature of international trade to identify intermediate inputs (Amiti et al., 2014; Brandt et al., 2012).

Data on firm-level characteristics are obtained from Micro.3, which includes census data on Italian firms with more than 20 employees from all sectors of the economy for the period 1989-2006. The database contains information on a number of variables appearing in a firm's balance sheet. For the purpose of this paper we use: number of employees, turnover, value added, capital, labour cost, intermediate inputs costs and capital assets. Capital is proxied by tangible fixed assets at book value (net of depreciation). Nominal variables are in million euros and are deflated using 2-digit industry-level production prices indices provided by ISTAT. After merging these two databases, we work with an unbalanced panel of about 48,179 manufacturing firms over the sample period.

In addition to firm-level data, we complement the analysis with information on country characteristics. We consider the two standard gravity-type variables, GDP_{jt} and $Distance_j$ to proxy for market size (Y_{jt}) and transportation costs (D_j), respectively. Data on GDP are taken from the World Bank's World Development Indicators database. Information on geographical distances are taken from CEPII and calculated following the great circle formula (De Sousa et al., 2012).

We augment the gravity model by including additional variables that might be expected to affect the costs of trading internationally. As predicted by equation (3) of our model, the probability of exporting depends on variable trade costs not related to distance (Δ_j), market specific fixed costs (F_j) and a multilateral resistance term (θ_{jt}). At the same time equation (4) suggests that a firm's export sales to a specific destination can be modelled in a parallel fashion to the model for export participation, though in this case market-specific fixed costs are not included.

For additional trade costs (Δ_j), we use a measure of average country-level import tariffs taken from the Fraser Institute ($Trade\ Opening_{jt}$)(Gwartney et al., 2014). This variable is a simple average of three sub-components: revenue from trade taxes, the mean tariff rate and the standard deviation of tariffs. Each sub-component is a standardized measure ranging from 0 to 10 which is increasing in the freedom to trade internationally.[14]

The market specific fixed costs (F_j) can be related to the establishment of a foreign distribution network, difficulties in enforcing contractual agreements, or the uncertainty of dealing with foreign bureaucracies. Following Bernard et al. (2015), to generate a proxy for these

[12]The database has been made available for work after careful screening to avoid disclosure of individual information. The data were accessed at the ISTAT facilities in Rome. The database has been built as a result of collaboration between ISTAT and a group of LEM researchers from the Scuola Superiore Sant'Anna, Pisa. See Grazzi et al. (2013) for further details.

[13]ISTAT collects data on trade based on transactions. A detailed description of requirements for data collection on trade is provided in Section 3.1 of the Technical Appendix. The section also provides additional descriptive statistics.

[14]As an robustness check, available upon request, we get the most-favored-nation tariffs (MFN tariffs) from the World Integrated Trade Solution (WITS) dataset.

costs we use information from three measures from the World Bank Doing Business dataset: number of documents for importing, cost of importing and time to import (Djankov et al., 2010). Given the high level of correlation between these variables, we use the primary factor ($Market\ Costs_j$) derived from principal component analysis as that factor accounts for most of the variance contained in the original indicators.

Finally, to proxy the multilateral resistance terms (θ_{jt}) we employ the variable $Remoteness_{jt}$ which captures the extent to which a country is separated from other potential trade partners. The idea is that a remote country has high shipping costs, high import prices, and thus a high aggregate price index. As in Manova and Zhang (2012) the variable remoteness is computed for each country as the distance weighted sum of the market sizes of all trading partners. Precisely, $Remoteness_j = \sum_{n=1}^{N} GDP_n * distance_{nj}$, where GDP_n is the GDP of the origin country and $distance_{nj}$ is the distance between n and j, and the summation is over all countries in the world n. Our results are robust to the use of an alternative measure of remoteness used in Baldwin and Harrigan (2011) given by $Remoteness_j = \sum_{n=1}^{N} (GDP_n/distance_{nj})^{-1}$.

4. Results

This section presents the results of our empirical analysis testing the main predictions of our theoretical model derived in section 2. We follow three steps. First, we provide evidence that importing has a positive effect on a firm's TFP. Second, we estimate the equation for a firm's export participation and for its export sales and show the influence that the component of TFP related to importing has on both the extensive and the intensive margin of exports. Third, we estimate the indirect impact that the two gravity forces have on a firm's exports due to the presence of imports in intermediates.

4.1. Imported intermediate inputs and firm productivity

Proposition 1 suggests that importing intermediate inputs increases a firm's productivity. Equation (5) derives an expression for a firm's TFP which depends on its initial productivity draw, (φ^f), the ratio of total intermediates over domestic inputs used, $\left(\frac{M^f_{tot}}{M^f_k}\right)$ and a set of variables which are constant at the firm-level.

As a first step of our empirical investigation, we estimate a firm's total factor productivity by using a gross output production function in the presence of input endogeneity (i.e. firms choose inputs based on their observed productivity level, the latter being unobserved by the econometrician) and by taking into account a firm's ability to import intermediates. We do that by relying on the method proposed by Gandhi et al. (2018) that we adapt to be consistent with our theoretical setting characterized by a Cobb-Douglas production function with a CES composite intermediate input. The proxy variable methods proposed by Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015) deal with the simultaneity (or transmission bias) problem mainly in the context of a value added production function (i.e. the intermediate input is not included in the estimated production function). However, Gandhi et al. (2017) has shown that a value-added production function can be constructed from an underlying gross output production function only under very restrictive hypotheses (such as the linear in intermediate inputs Leontief specification) that are not compatible with our theoretical framework. Given that proxy variable methods are likely to suffer from identification issues when employed with a gross output production function,[15] we follow the suggestions of Ackerberg et al. (2015) and Gandhi et al. (2018) by exploiting the information contained in the first order condition of a firm's static profit maximization problem with respect to intermediate inputs.[16]

[15]On this point see Bond and Söderbom (2005), Ackerberg et al. (2015) and Gandhi et al. (2018).
[16]Section 3.2 of the Technical Appendix provides details of the TFP estimation.

Table 1: Production function estimates

	$\ln M^f_{tot,t}$	$\ln l^f_t$	$\ln k^f_t$	$\ln \frac{M^f_{tot,t}}{M^f_t}$	N.Obs
	(1)	(2)	(3)	(4)	(5)
Food, Beverages and Tobacco	0.70***	0.20***	0.07***	0.26***	9875
Textiles and Apparel	0.49***	0.36***	0.07***	0.54***	16570
Hide and Leather	0.62***	0.31***	0.06***	0.59***	6517
Wood and Cork	0.65***	0.28***	0.06***	0.31***	3751
Pulp and Paper	0.73***	0.22***	0.04***	0.24***	3350
Printing and Publishing	0.59***	0.36***	0.04***	0.37***	4847
Coke and Chemical products	0.73***	0.24***	0.02***	0.22***	6334
Rubber and Plastics	0.68***	0.25***	0.04***	0.29***	9258
Processing of non-metallic minerals	0.66***	0.24***	0.09***	0.28***	8381
Basic Metals	0.71***	0.28***	0.039***	0.22***	4268
Fabricated Metal Products	0.56***	0.30***	0.07***	0.53***	23745
Machinery and Equipment	0.63***	0.37***	0.02***	0.29***	21647
Electrical and Optical Equipment	0.60***	0.36***	0.04***	0.32***	12297
Motor Vehicles and Trailers	0.65***	0.32***	0.04***	0.17***	2855
Other Transport Equipment	0.53***	0.39***	0.13***	0.36***	1726
Other manufacturing industries	0.68***	0.27***	0.06***	0.17***	9876

Note: The table reports production function estimates by sector using data on 2000-2006. Column (1) reports the coefficient of intermediate inputs ($M^f_{tot,t}$), column (2) the coefficient of labour (l^f_t), column (3) the coefficient of capital (k^f_t) and column (4) the coefficient of the ratio of intermediate inputs on domestic inputs ($M^f_{tot,t}/M^f_t$). Asterisks denote significance levels obtained with bootstrapped standard errors (500 replications) (***:p<1%; **: p<5%; *: p<10%).

For each sector h, we consider the production function used in our theoretical model augmented with physical capital[17]

$$\ln y^f_t = \beta_0 + \alpha \ln M^f_{tot,t} + \beta_l \ln l^f_t + \beta_k \ln k^f_t + \frac{\alpha}{\phi - 1} \ln \frac{M^f_{tot,t}}{M^f_t} + \ln \varphi^f_t + \epsilon^f_t \tag{10}$$

where y^f_t is the sales of firm f at time t, $M^f_{tot,t}$ is total intermediate inputs, l^f_t is labor, k^f_t stands for the capital stock, and M^f_t corresponds to domestic intermediate inputs. The error can be decomposed into a productivity shock φ^f_t, observable to firms but not to the econometrician, and an i.i.d. component ϵ^f_t. The constant, β_0, subsumes common industry-level factors.[18]

Following the proxy variable methods and Gandhi et al. (2018), the law of motion of $\ln \omega^f_t$ is represented by a first order Markov process. In order to take into account the possibility of the existence of an endogenous productivity process characterized by a dynamic learning by importing effect, we follow Kasahara and Rodrigue (2008) and, for the main results of the paper, we let the dynamics of productivity to potentially depend on past importing behavior.[19]

Table 1 presents the results of the production function estimates. The estimated coefficients for the ratio of total over domestic intermediate inputs in equation (10) are always positive and statistically significant across different sectors, pointing to the importance of foreign intermediates in explaining productivity differences across firms within sectors. At one extreme, for the "Textiles and Apparel", "Hide and Leather" and "Fabricated Metal Products" sectors, we find

[17] To simplify the notation, in the estimation equation we omit the subscript k when referring to the domestic country, that in our case is only Italy.

[18] Such as $\left(\frac{\beta_{mk} Y_k}{Y}\right)^{\frac{\alpha_h}{\phi_h - 1}} \left(\frac{(1+\pi)}{Y}\right)^{\frac{\alpha_h}{1-\phi_h}}$ in equation (5).

[19] In section 4.2.1 we let the dynamics of productivity to potentially depend on importing and exporting behavior, therefore also allowing for the possibility of the existence of learning by exporting effects as in De Loecker (2013). This is shown in equations (17), (18) and (19) of the Technical Appendix. Table 6 of the Technical Appendix reports the estimated learning by importing and learning by exporting effects by sector. The results suggest that the learning by importing effects are positive and significant, while the learning by exporting effects are of much lower magnitude and statistically significant only for about half of the sectors.

that a 10% rise in the ratio of intermediate inputs on domestic inputs would increase productivity by 5.3% to 5.9%. At the bottom of the sectoral distribution, this effect amounts to 1.7% for the "Motor Vehicles and Trailers" and "Other manufacturing Industries" sectors.[20]

4.2. The extensive and intensive margins of exports

Equations (3) and (4) describe how a firm's decision to export and its export value to a country are related to gravity forces both through a direct effect and an indirect effect due to the TFP contribution of trade in intermediates. These two equations form the underpinning of our estimations. Therefore, a model for a firm's decision to export to a specific country can be specified as follows

$$ExportStatus^f_{jt} = b_0 + b_1 \ln \widehat{\varphi}^f_t + b_2 \ln \widehat{\chi}^f_t + b_3 \ln D_j + b_4 \ln Y_{jt} + b_5 \Delta_{jt} + b_6 F_j + b_7 \ln \theta_{jt} + d^f + d_i + \epsilon^f_{jt} \tag{11}$$

where the dependent variable, $ExportStatus^f_{jt}$, is a dummy variable that takes value one if a firm f exports to country j at time t and zero otherwise. The empirical specification includes our estimates for a firm's innate productivity, $\widehat{\varphi}^f_t$, and for the TFP-enhancing effect of imported intermediate inputs, $\ln \widehat{\chi}^f_t = \widehat{\frac{\alpha}{\phi-1}} \ln \frac{M^f_{tot,t}}{M^f_t}$. In accordance with our model we expect both b_1 and b_2 to be positive. In addition, the equation includes all the country-level variables that appear in equation (3) (Y_{jt}, θ_{jt}, D_j, Δ_{jt}, F_j). The model predicts that the probability of serving the foreign market j increases with the size of the country ($b_4 > 0$) and the level of remoteness ($b_7 > 0$), while it decreases with the level of variable costs ($b_3 < 0$; $b_5 < 0$) and fixed costs ($b_6 < 0$).

Following Bernard and Jensen (2004), to estimate our binary choice framework with unobserved heterogeneity, we employ a linear probability model so that firm fixed-effects are accounted for in the regressions. Although this estimation strategy suffers from the problem of predicted probabilities outside the 0-1 range, it allows us to control for any unobserved time constant firm characteristic that influences the decisions regarding entry into foreign markets. By exploiting the three-dimensional nature (firms, destinations, time) of our dataset, we include firm fixed-effects (d^f) to account for time-invariant firm-level unobserved heterogeneity. Moreover, we introduce year-geographical areas dummies (d_i) to account for all the time-variant shocks common to countries belonging to the same area. We group countries in 20 different areas, as done in Serti and Tomasi (2014). Standard errors are clustered at the firm and destination-level.

We next explore whether firm and country differences are relevant for determining how much a firm sells across different markets, that is the intensive margin of exports. The econometric model, which can be thought of as a micro-gravity equation, takes the following form

$$\ln Exports^f_{jt} = c_0 + c_1 \ln \widehat{\varphi}^f_t + c_2 \ln \widehat{\chi}^f_t + c_3 \ln D_j + c_4 \ln Y_{jt} + c_5 \Delta_{jt} + c_6 \ln \theta_{jt} + d^f + d_i + \epsilon^f_{jt} \tag{12}$$

where the dependent variable is the (log) total exports of a firm f to country j at time t. As in the previous equation, we include a firm's innate productivity, the TFP component related to the use of imported inputs, and country determinants (Y_{jt}, θ_{jt}, D_j, Δ_{jt}). Following equation (4), we exclude the trade fixed costs variable. As for the export decision equation, we run the regression controlling for firm and year-area fixed-effects and clustering the errors at the firm and destination-level.

[20]Comparing the estimated coefficients for $\ln M^f_{tot,t}$ and $\ln \frac{M^f_{tot,t}}{M^f_t}$ we recover an average estimated elasticity of substitution for intermediate inputs ($\hat{\phi}$) of about 3.3, which is very similar to what found by Kasahara and Rodrigue (2008).

To take into account firms' unobserved heterogeneity, we estimate equation 12 also using the level of exports as dependent variable by employing a conditional (firm) fixed-effects Poisson model, which is appropriate for nonlinear models such as the gravity equation (Silva and Tenreyro, 2006). The main advantage of the Poisson estimator is that it naturally includes observations for which the observed trade value is zero, that is it takes into account the extensive and the intensive margins at the same time. Such observations are dropped from the OLS model because the logarithm of zero is undefined. However, especially at the firm-level, zero trade flows are very common, since not all firms are trading with all partners.[21]

One of the main problems in estimating equations (11) and (12) concerns the potential endogeneity of our key covariate, that is the estimated TFP-enhancing effect of imported intermediate inputs ($\widehat{\chi}_t^f$), which is a positive function of the share of imported intermediate inputs. The introduction of firm fixed-effects ensures that our results are not driven by time constant unobserved heterogeneity which is correlated with the imported inputs decisions. However, endogeneity can arise because of time variant omitted variables, simultaneity problems, or measurement error.

First, in estimating the evolution of $\widehat{\varphi}_t^f$, we incorporate possible learning by importing effects and we rely on lagged inputs sourcing strategies in the moment conditions. This should reduce the likelihood that the error term contains unobserved productivity shocks that affect both our key variable and a firm's sales (abroad and at home). However, we cannot rule out that a firm changes its share of imported intermediate inputs as a reaction to cost and/or demand shocks which are not picked up by $\widehat{\varphi}_t^f$.[22] A positive correlation between these productivity shocks and the relative use of imported intermediates would induce an upward bias in the estimates of the $\widehat{\chi}_t^f$ coefficient. Second, although our paper focuses on the causal effect of importing intermediates on exporting, in the presence of learning-by-exporting causality may run also in the other direction: by expanding their exports firms become more efficient and, as a consequence, increase their use of imported intermediate inputs. This would make the estimates of $\widehat{\chi}_t^f$ coefficient be downward biased. Third, our main independent variable is likely to be measured with error because we cannot observe the prices of intermediate inputs and we are using aggregate sectoral prices to deflate their observed values. If expansions in the deflated value of domestic (imported) intermediate inputs understates (overstates) the actual increase in the use of domestic (imported) intermediate inputs, the coefficient for $\widehat{\chi}_t^f$ would be downward biased.

To identify the causal effect of the TFP related to imported inputs on firms' export activities we apply an instrumental variable approach.[23] Precisely, we use two instrumental variables. Following previous work by Mion and Zhu (2013), we construct a firm-level instrument based on the gross domestic product (GDP) of a country[24] that proxies for the number of available varieties of foreign intermediate inputs. Starting from this macro variable, we compute a firm-level instrument by taking a weighted average where the weights reflect the relative importance of the different source countries in a firm's total imported inputs. Specifically, we construct, for each firm, the weighted average of $\ln GDP$, denoted as IV_{GDP_f}, using as weights a firm's import share of each country. In order to address issues related to changes, across products or countries, in a firm's imported input mix due to variations in this macro variable, we rely on constant weights computed as the import shares of the initial year. As a robustness check, we adopt alternative weighting strategies, e.g., one year lagged import shares or the import shares

[21] As an additional robustness check, to control for possible selection bias we also employ a two-stage procedure in the spirit of Heckman's method by including the polynomials of the predicted value of $ExportStatus_{jt}^f$, obtained after estimating equation (11), into equation (12). This alternative specification is shown in Table 11 of the Technical Appendix.

[22] Since we are using a revenue production function, $\widehat{\varphi}_t^f$ contains both efficiency and demand factors.

[23] Table 7 of the Technical Appendix reports the estimation results of equations (11) and (12) in the paper, by using a simple OLS approach.

[24] The information on GDP comes from the World Bank database.

Table 2: Firms' exports extensive and intensive margin by country: instrumental variables

Dep. Var.	$ExportStatus^f_{jt}$	$\ln Export^f_{jt}$	$Export^f_{jt}$	$ExportStatus^f_{jt}$	$\ln Export^f_{jt}$	$Export^f_{jt}$	$ExportStatus^f_{jt}$	$\ln Export^f_{jt}$	$Export^f_{jt}$
	OLS		Poisson	OLS		Poisson	OLS		Poisson
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	IV_{GDP_f}			$IV^P_{EUM_f}$			$IV_{GDP_f} IV^P_{EUM_f}$		
$\ln \widehat{\varphi}^f_t$	0.108***	1.152***	2.368***	0.102***	1.287***	2.118***	0.105***	1.187***	2.198***
	(0.016)	(0.146)	(0.389)	(0.016)	(0.197)	(0.386)	(0.015)	(0.153)	(0.329)
$\ln \widehat{\chi}^f_t$	0.604***	5.000***	11.962***	0.553***	6.234***	9.745***	0.571***	5.341***	10.263***
	(0.084)	(0.987)	(3.111)	(0.085)	(1.428)	(2.693)	(0.076)	(1.021)	(2.301)
$\ln GDP_{jt}$	0.057***	0.491***	0.834***	0.057***	0.492***	0.810***	0.058***	0.494***	0.834***
	(0.005)	(0.027)	(0.018)	(0.005)	(0.027)	(0.024)	(0.005)	(0.027)	(0.018)
$\ln Distance_j$	-0.083***	-0.506***	-0.936***	-0.084***	-0.505***	-0.992***	-0.085***	-0.509***	-0.940***
	(0.013)	(0.057)	(0.058)	(0.013)	(0.057)	(0.054)	(0.013)	(0.057)	(0.059)
$Trade\ Opening_{jt}$	0.010***	0.049*	0.044**	0.010**	0.049*	0.096**	0.010**	0.050*	0.045***
	(0.004)	(0.026)	(0.017)	(0.004)	(0.026)	(0.018)	(0.004)	(0.026)	(0.017)
$\ln Remoteness_{jt}$	0.085*	0.427*	1.017***	0.085*	0.421*	0.445***	0.086*	0.424*	1.042***
	(0.044)	(0.217)	(0.212)	(0.044)	(0.216)	(0.172)	(0.044)	(0.217)	(0.211)
$Market\ Costs_j$	-0.014**		-0.298***	-0.014**		-0.146***	-0.015**		-0.296***
	(0.006)		(0.060)	(0.006)		(0.050)	(0.006)		(0.061)
Year*Area FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8,582,803	1,435,557	8,582,803	8,504,615	1,426,550	8,504,615	8,225,922	1,389,028	8,225,922
adj. R^2	0.348	0.333		0.349	0.332		0.350	0.334	
Underidentification stat.	84.896	38.937		81.884	37.693		91.940	41.767	
(p-value)	0.000	0.000		0.000	0.000		0.000	0.000	
Weak identification stat.	262.532	183.135		229.811	134.748		171.375	115.878	
Hansen J stat.							0.451	1.216	1.017
(p-value)							0.502	0.270	0.313

Note: The table reports regressions using data on 2000-2006 for the extensive margin (columns 1-4-7), the intensive margin (2-5-8), and the Poisson model (columns 3-6-9). In Columns 1-3 the instrumental variable IV_{GDP_f} is built by weithing ln GDP by a firm's import share of each country in the initial year. In Columns 4-6 the instrumental variable $IV^P_{EUM_f}$ is built by weighting the total imports of European countries by the relative importance of a product in a firm's total imports in the initial year. In columns 7-9 the instrumental variables used are IV_{GDP_f} and $IV^P_{EUM_f}$ together. Robust standard errors clustered at the firm and destination-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

of the first year in which a firm is observed importing. By employing this instrument we aim at exploiting the exogenous variation of this source weighted macro characteristic as predictor of changes in the usage of imported intermediate inputs at the firm-level.

As a second instrument, similarly to Hummels et al. (2014), we build a variable based on the total imports of other European countries, excluding Italy. The total European countries imports are obtained from the COMTRADE dataset at the level of HS6 products. We create a firm-level index ($IV^P_{EUM_f}$) using as weights the relative importance of a product in a firm's total imports during the initial year. This IV strategy exploits the variation of aggregate imports at the product-level for similar developed countries as predictor of changes in the TFP enhancing effect of firm-level imports. The exclusion restriction is based on the hypothesis that aggregate import dynamics at the product-level for other European countries are mainly determined by supply-side cost and technology factors which are sufficiently exogenous to a firm's export performance. However, it is possible that the increase in imports of a particular intermediate input is determined by an increase in the international demand of the corresponding final products. Our results are robust to controlling for this by using the total exports of other European countries, weighted by the relative importance of a product in a firm's total exports in the initial year. This variable will reflect international demand shocks for a firm's exported products.[25]

[25] Similar results are obtained if we use an alternative instrument the total imports from China of other European countries and, as before, weighting these trade flows by using the relative importance of a product in a firm's total imports during the initial year. The main insight of these IVs is to exploit the increase in Chinese exports to other developed countries as an exogenous supply shock. Indeed, during our sample period

Table 3: Firms' exports extensive margin by country: IV. Robustness Checks

Dep. Var.	$ExportStatus^f_{jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
	Learning by exporting	Excluding Only-Exporters	Excluding Importers of No-Intermediates	Firm-Destination and Year-Destination FE		
					No Imports Dest.	No Imports Dest Learning by Exporting
$\ln \widehat{\varphi}^f_t$	0.104***	0.094***	0.131***	0.113***	0.094***	0.093***
	(0.015)	(0.014)	(0.020)	(0.013)	(0.013)	(0.013)
$\ln \widehat{\chi}^f_t$	0.562***	0.605***	0.733***	0.623***	0.504***	0.498***
	(0.075)	(0.081)	(0.108)	(0.059)	(0.066)	(0.065)
$\ln GDP_{jt}$	0.058***	0.058***	0.064***			
	(0.005)	(0.005)	(0.006)			
$\ln Distance_j$	-0.085***	-0.085***	-0.090***			
	(0.013)	(0.013)	(0.014)			
$Trade\ Opening_{jt}$	0.010***	0.010**	0.011**			
	(0.004)	(0.004)	(0.004)			
$\ln Remoteness_{jt}$	0.088**	0.088**	0.092*			
	(0.044)	(0.044)	(0.046)			
$Market\ Costs_j$	-0.015**	-0.015**	-0.017**			
	(0.006)	(0.006)	(0.007)			
Year*Area FE	Yes	Yes	Yes	-	-	-
Firm FE	Yes	Yes	Yes	-	-	-
Firm-Destination FE	-	-	-	Yes	Yes	Yes
Year-Destination FE	-	-	-	Yes	Yes	Yes
N	8,217,355	8,179,448	6,276,792	7,940,616	7,401,092	7,392,503
adj. R^2	0.350	0.350	0.363	0.791	0.765	0.765
Underidentification stat.	92.011	90.811	75.982	322.398	88.019	88.069
(p-value)	0.000	0.000	0.000	0.000	0.000	0.000
Weak identification stat.	172.126	163.303	96.847	172.809	172.520	173.214
Hansen J stat.	0.375	0.423	0.558	0.242	0.135	0.098
(p-value)	0.540	0.515	0.455	0.623	0.713	0.754

Note: The table reports regressions for the extensive margin using data on 2000-2006 and employing as instrumental variables IV_{GDP_f}, which is built by weighting ln GDP by a firm's import share of each country in the initial year, and $IV^P_{EUM_f}$, which is built by weighting the total imports of European countries by the relative importance of a product in a firm's total imports in the initial year. In columns 1 and 6 we re-estimate the two TFP components by allowing the law of motion of φ^f_t to endogenously depend on the export share. In column 2 we exclude the only-exporters from the analysis. In column 3 we consider only those firms importing intermediate inputs. In columns 4-6 we control for firm-destination and year-destination fixed-effects. In columns 5 and 6 we exclude from the sample the export destinations from which a firm is importing. Robust standard errors clustered at the firm and destination country-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

Estimation results for the extensive margins using as instrument IV_{GDP_f}, $IV^P_{EUM_f}$ and both of them are shown in columns 1,4,7 of Table 2, respectively. For the intensive margin, columns 2,5,8 of Table 2 report the coefficients using the linear regression model. Finally, columns 3,6,9 show the Poisson specification.[26] At the bottom of the Table we report the under-identification (Kleibergen-Paap LM), weak identification (Kleibergen-Paap Wald F), and over-identifying restriction (Hansen J) statistics and p-values. The statistics of the first two tests indicate that our instruments have predictive power and the Hansen test suggests that our instruments are valid.[27]

The results provide a clear picture. Column 1 shows that a firm's probability of exporting to a destination is positively affected by both the innate productivity ($\widehat{\varphi}^f_t$) and the TFP-enhancing effect of imported intermediate inputs ($\widehat{\chi}^f_t$). A 10 percent increase in the innate productivity is associated with an increase of about 1.1 percentage points in the probability of exporting

there has been an impressive increase of imports from China, mainly due to its growth in competitiveness and its accession to the WTO. A similar identification strategy is used by Autor et al. (2013); Donoso et al. (2015); Dauth et al. (2018).

[26] For both the linear and the Poisson specification we use a GMM estimator. For the former we use the moment conditions described in Blundell et al. (2002) and Agrawal et al. (2014): $E\left[\left(y_{it} - \mu_{it}\frac{\overline{y_i}}{\overline{\mu_i}}\right) z_{it}\right] = 0$, where $\mu_{it} = \exp(x_{it}\beta)$ and $(\overline{y_i}, \overline{\mu_i})$ are means of the outcome and the predicted outcomes at the firm-level.

[27] Table 8 of the Technical Appendix reports the estimates of the first stage of the IV estimation for all the equations of Table 2, suggesting that the IVs are, as expected, positively correlated with $\ln \widehat{\varphi}^f_t$.

to a destination. The magnitude of this effect is sizable if compared with the probability of exporting to a country observed in our sample, which is about 7.5 percent. This means that a firm's probability of exporting to a country rises of approximately 14 percent, following a 10 percent increase in $\widehat{\varphi}_t^f$. The coefficient for the contribution of imported intermediate inputs to TFP is positive and statistically significant. The coefficient is higher in magnitude than that observed for the innate productivity: a 10 percent increase in $\widehat{\chi}_t^f$ is associated with an increase of about 6 percentage points in the probability of exporting to a destination. The result provides evidence for the relevance of the TFP-enhancing effect of imported intermediates: a rise of 10 percent of $\widehat{\chi}_t^f$ increases the probability of exporting by more than 75 percent.

As for the two gravity variables, we find that the probability of exporting to a specific market increases with market size but decreases with distance. A 10 percent rise in the destination country's GDP is associated with an increase of 0.6 percentage points in the probability of exporting to that country. A 10 percent increase in distance decreases the likelihood of a positive export decision by approximately 0.8 percentage points. As above, to gauge the economic significance of these variables we compare the estimated effects with the observed probability of exporting. The coefficient for market size suggests that, holding all other independent variables constant, a 10 percent increase in the GDP of a country raises the probability of exporting to that market by about 8 percent. The *ceteris paribus* effect of a 10 percent increase in distance is a decrease in the probability of exporting of around 10 percent.

Concerning the other country properties, as expected the probability of exporting decreases with market costs. The negative and significant coefficient of $Market\ Costs$ suggests the existence of country-specific fixed export costs: the lower these costs are, the higher the probability of reaching a market. Easy and accessible markets are likely to be served by a large number of firms, whereas less accessible countries with higher fixed export costs are more difficult to export to. The coefficients for *Remoteness* and $Trade\ Opening$ have both the expected positive sign. Since remoteness makes a destination market less competitive, *ceteris paribus*, it is relatively easier for a firm to serve a trade partner that is geographically isolated from most other nations. The probability of exporting to a country should indeed increase with both the remoteness of the destination and its level of freedom to trade. As shown in columns 4 and 7 of Table 2, our findings are robust to using as instrument the total imports of other European countries (column 4), and when combining both instruments together the Hansen test suggests that our instruments are valid (column 7).

Column 2 of Table 2 reports the result for the intensive margin using the linear regression model. The estimated parameters display the expected signs. We confirm that both the innate productivity and the TFP-enhancing effect of imported intermediate inputs positively affect a firm's exports to a country. More productive firms export more to each country: a 10 percent increase in a firm's innate productivity increases its exports by approximately 12 percent. Even stronger is the effect of productivity due to imported intermediate inputs: exports increase by approximately 50 percent following a rise of 10 percentage in $\widehat{\chi}_t^f$. The estimated elasticities of exports to GDP and $Distance$ are 0.49 and -0.51, respectively. These results are directly comparable with the previous work by Lawless and Whelan (2014) that use a panel survey of Irish firms to provide empirical support for the heterogeneous-firm models for firm-level patterns of trade across destinations. Their estimated elasticities are 0.58 and -0.52, respectively, consistent with our figures.[28] Finally, the estimated effects of *Remoteness* and $Trade\ Opening$ show the expected positive signs and are statistically significant. As before, columns 5 and 8 confirm that our findings are robust to the use of the other instrument and the Hansen test suggests that our instruments are valid.

Columns 3, 6 and 9 consider the estimation of Equation (12) in its multiplicative form with a pseudo-maximum-likelihood technique. Looking at the results we can conclude that the main

[28]While Lawless and Whelan (2014) provide evidence also for the extensive margin, their results can not be directly compared with ours as they report the probit coefficients rather than the marginal effects.

Table 4: Firms' exports intensive margin by country: IV. Robustness Checks

Dep. Var.	$\ln Exports^f_{jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
	Learning by exporting	Excluding Only-Exporters	Excluding Importers of No-Intermediates	Firm-Destination and Year-Destination FE		
					No Imports Dest.	No Imports Dest Learning by Exporting
$\ln \widehat{\varphi}^f_t$	1.153***	1.043***	1.300***	1.847***	1.492***	1.490***
	(0.146)	(0.137)	(0.202)	(0.179)	(0.167)	(0.165)
$\ln \widehat{\chi}^f_t$	4.944***	5.386***	5.976***	10.040***	7.882***	7.709***
	(0.960)	(1.070)	(1.348)	(1.076)	(1.145)	(1.111)
$\ln GDP_{jt}$	0.496***	0.496***	0.511***			
	(0.027)	(0.027)	(0.027)			
$\ln Distance_j$	-0.508***	-0.509***	-0.524***			
	(0.057)	(0.057)	(0.059)			
$Trade\ Opening_{jt}$	0.048*	0.048*	0.050*			
	(0.026)	(0.026)	(0.027)			
$\ln Remoteness_{jt}$	0.424*	0.424*	0.453**			
	(0.217)	(0.217)	(0.224)			
Year*Area FE	Yes	Yes	Yes	-	-	-
Firm FE	Yes	Yes	Yes	-	-	-
Firm-Destination FE	-	-	-	Yes	Yes	Yes
Year-Destination FE	-	-	-	Yes	Yes	Yes
N	1,388,023	1,384,598	1,192,005	1,256,783	906,619	905,991
adj. R^2	0.334	0.334	0.336	0.820	0.789	0.789
Underidentification stat.	41.803	41.381	38.738	37.222	45.130	45.178
(p-value)	0.000	0.000	0.000	0.000	0.000	0.000
Weak identification stat.	116.713	110.970	67.892	109.478	101.582	102.936
Hansen J stat.	1.242	1.480	0.893	0.360	0.018	0.028
(p-value)	0.265	0.224	0.345	0.548	0.892	0.866

Note: The table reports regressions for the intensive margin using data on 2000-2006 and employing as instrumental variables IV_{GDP_f}, which is built by weighting ln GDP by a firm's import share of each country in the initial year, and $IV^P_{EUM_f}$, which is built by weighting the total imports of European countries by the relative importance of a product in a firm's total imports in the initial year. In columns 1 and 6 we re-estimate the two TFP components by allowing the law of motion of φ^f_t to endogenously depend on the export share. In column 2 we exclude the only-exporters from the analysis. In column 3 we consider only those firms importing intermediate inputs. In columns 4-6 we control for firm-destination and year-destination fixed-effects. In columns 5 and 6 we exclude from the sample the export destinations from which a firm is importing. Robust standard errors clustered at the firm and destination country-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

message with respect to the previous specifications does not change. The estimated elasticity of exports with respect to both the innate productivity and the TFP component related to importing is economically and statistically significant. The interpretation of the coefficients from the Poisson model is straightforward, and follows exactly the same pattern as under OLS, that is the coefficients of any independent variables entered in logarithms can be interpreted as simple elasticities.

4.2.1. Robustness checks

In this section, we consider a set of exercises aimed at testing the robustness of our results to alternative estimates of a firm's TFP, to changes in the sample composition, and to the adoption of alternative fixed-effects. These robustness checks are reported in Table 3 for the extensive margin, Table 4 for the intensive margin, and Table 5 for the Poisson model, using as instruments both the IV_{GDP_f} and $IV^P_{EUM_f}$.

The existence of learning by exporting effects could create reverse causality problems which we tried to address by using IVs. However, there is also the possibility that the variable $\hat{\chi}^f_t$ mechanically contains learning by exporting effects because the import share of a firm is positively correlated to its export share. Therefore, we have re-estimated the two TFP components by allowing the law of motion of φ^f_t to endogenously depend on the export share.[29] Results obtained by using this alternative TFP estimation strategy are reported in columns 1

[29] Table 6 of the Technical Appendix shows the coefficients for the learning by exporting effects by sector.

Table 5: Firms' exports by country (Poisson): IV. Robustness Checks

Dep. Var.	$Exports^f_{jt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
	Learning by exporting	Excluding Only-Exporters	Excluding Importers of No-Intermediates	Firm-Destination and Year-Destination FE	No Imports Dest.	No Imports Dest Learning by Exporting
$\ln \widehat{\varphi}^f_t$	2.246***	2.116***	3.104***	2.408***	2.204***	2.159***
	(0.326)	(0.321)	(0.663)	(0.271)	(0.211)	(0.193)
$\ln \widehat{\chi}^f_t$	11.137***	12.085***	14.799***	11.820***	10.268***	10.337***
	(2.426)	(2.705)	(5.051)	(1.879)	(1.997)	(2.027)
$\ln GDP_{jt}$	0.772***	0.773***	0.773***			
	(0.022)	(0.022)	(0.022)			
$\ln Distance_j$	-0.962***	-0.965***	-0.961***			
	(0.052)	(0.052)	(0.055)			
$Trade\ Opening_{jt}$	0.089***	0.090***	0.088***			
	(0.017)	(0.017)	(0.018)			
$\ln Remoteness_{jt}$	0.389**	0.400**	0.404*			
	(0.178)	(0.179)	(0.188)			
$Market\ Costs_j$	-0.139***	-0.135***	-0.133**			
	(0.052)	(0.052)	(0.055)			
Year*Area FE	Yes	Yes	Yes	-	-	-
Firm FE	Yes	Yes	Yes	-	-	-
Firm-Destination FE	-	-	-	Yes	Yes	Yes
Year FE	-	-	-	Yes	Yes	Yes
N	8,217,355	8,179,448	6,276,792	7,940,616	7,401,092	7,392,503
Hansen J stat.	0.257	0.346	0.541	1.339	0.189	0.220
(p-value)	0.6119	0.557	0.462	0.247	0.664	0.639

Note: The table reports regressions for the Poisson model using data on 2000-2006 and employing as instrumental variables IV_{GDP_f}, which is built by weighting ln GDP by a firm's import share of each country in the initial year, and $IV^P_{EUM_f}$, which is built by weighting the total imports of European countries by the relative importance of a product in a firm's total imports in the initial year. In columns 1 and 6 we re-estimate the two TFP components by allowing the law of motion of φ^f_t to endogenously depend on the export share. In column 2 we exclude the only-exporters from the analysis. In column 3 we consider only those firms importing intermediate inputs. In columns 4-6 we control for firm-destination and year-destination fixed-effects. In columns 5 and 6 we exclude from the sample the export destinations from which a firm is importing. Robust standard errors clustered at the firm-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

and 6 of Table 3 for the extensive margin, columns 1 and 6 of Table 4 for the intensive margin, and columns 1 and 6 of Table 5 for the Poisson specification. The results are not affected by taking directly into account possible learning by exporting mechanisms in the TFP estimation: the $\hat{\chi}^f_t$ is still positive and statistically significant, and the point estimate is statistically equal (within 1 standard error band) to the baseline estimates.

Next, we re-estimate the baseline specification on different sub-samples to verify that our main results do not crucially depend on the peculiar behavior of specific groups of firms. First, we exclude from the analysis those firms that are only exporters. This paper focus on equilibria where a firms engaged in international trade are either two-way traders or just-only importers. Our data confirms that the majority of firms are involved in both trade activities while only a small fraction exports without importing. Column 2 of Tables 3, 4 and 5 presents the results by dropping the only exporters. Second, in column 3 of the three Tables, we exclude from the analysis those firms that source from abroad both capital and intermediate goods and consider only those firms importing intermediate inputs, defined as those falling into the intermediate input category according to BEC classification system. The findings are robust to these changes in the sample coverage, which affect neither the sign of the coefficients nor their significance.

As an alternative specification, in columns 4-6 of Tables 3, 4 and 5, we estimate the equations including firm-destination and year-destination fixed-effects. In this case, identification of our key variable, $\hat{\chi}^f_t$, relies only on variations over time of a firm's exports to the same destination, controlling for time variant and time invariant country characteristics. In our theoretical framework the effect of importing intermediate inputs comes only through $\hat{\chi}^f_t$. However, besides the TFP mechanism, there could be additional channels through which importing intermediate inputs influences exporting. In particular, one could imagine that importing from country j

Table 6: Import ratio elasticities

Dep. Var.	$\ln \frac{M^f_{jt}}{M^f_t}$		$\frac{M^f_{jt}}{M^f_t}$	
	(1)	(2)	(3)	(4)
$\ln GDP_{jt}$	0.167***	0.167***	0.822***	0.825***
	(0.009)	(0.009)	(0.027)	(0.026)
$\ln Distance_j$	-0.186***	-0.186***	-0.796***	-0.836***
	(0.015)	(0.015)	(0.089)	(0.072)
$\frac{Export^f_{jt}}{DomesticSales^f_t}$		0.001		0.001***
		(0.001)		(0.000)
$Market\ Costs_j$			-0.046	-0.008
			(0.058)	(0.043)
Year*Area FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
N	493,275	493,275	17,987,780	17,987,780
adj. R^2	0.326	0.326		

Note: The table reports the results of the OLS (columns 1-2) and of the Poisson Pseudo Maximum Likelihood (columns 3-4) estimators using data on 2000-2006. The dependent variable is the import ratio (M^f_{jt}/M^f_t) in logarithm (columns 1-2) and in value (columns 3-4). All the regressions include a constant term. Robust standard errors clustered at the firm and destination-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

reduces the fixed or variable cost of exporting to country j. By controlling for firm-destination and year-destination fixed-effects we can reduce this issue to the extent it is connected to time-constant unobserved heterogeneity at the firm-country level or to unobserved country-level determinants of trade flows which are common at the import and at the export side.[30]

In addition, in both columns 5 and 6 we exclude from the sample the export destinations from which a firm is importing. In this way possible market specific cost externalities from importing to exporting are shut down. If these factors were driving our results we would expect a decrease in the estimated effect of $\hat{\chi}^f_t$. The estimation results confirm our main findings.

4.3. The indirect effect of gravity forces

The aim of this section is to quantify the indirect impact of gravity forces on a firm's export behaviour through importing. As indicated by equations (6)-(9), to do that we first need to compute the elasticity of χ_{hk} with respect to the two gravity forces. Then, we have to multiply the elasticity of χ_{hk} with respect to either distance or market size by the elasticity of exports to χ_{hk}, obtained through the export gravity equations. In this way we obtain the elasticity of exports to distance and market size through importing.

Let's start with the computation of the elasticity of χ_{hk} with respect to distance from country j, $\rho^f_{D_j}$, which can be written as[31]

$$\rho^f_{D_j} = \frac{d\ln \chi_{hk}}{d\ln D_{kj}} = \frac{\alpha_h}{\phi_h - 1} * \frac{M^f_j}{\sum_{n=1}^{N} M^f_n} * \frac{d\ln\left(\frac{M^f_j}{M^f_k}\right)}{d\ln D_{kj}}. \tag{13}$$

Similarly, the elasticity of χ_{hk} with respect to market size of country j, $\rho^f_{Y_j}$, is given by

[30] Alternatively, we estimate the coefficients of $ln\varphi^f_t$ and $ln\chi^f_t$ by including year-destination and firm-destination fixed effects and then regress the estimated fixed effects on the country-level variables (Y_{jt}, θ_{jt}, D_j, Δ_{jt}, F_j) that appear in equations (11) and (12). The results are reported in Tables 9 and 10 of the Technical Appendix. For both the extensive and the intensive margins, the magnitude of the estimated coefficients of our country-level variables are comparable with those obtained in Table 2 of the paper.

[31] The derivation of equations 13 and 14 can be found in Section 2.4 of the Technical Appendix.

Table 7: Average indirect effects of gravity forces on export margins: by origin of imports

Import-Source Country	Extensive Margin Ψ_{Y_j} (1)	Ψ_{D_j} (2)	Intensive Margin Ψ_{Y_j} (3)	Ψ_{D_j} (4)	Both Margins Ψ_{Y_j} (5)	Ψ_{D_j} (6)
Germany	0.004	-0.004	0.034	-0.033	0.066	-0.064
France	0.002	-0.002	0.017	-0.016	0.032	-0.031
China	0.001	-0.001	0.009	-0.008	0.017	-0.016
Spain	0.001	-0.001	0.008	-0.008	0.015	-0.015
Belgium	0.001	-0.001	0.008	-0.008	0.015	-0.015
Austria	0.001	-0.001	0.008	-0.008	0.015	-0.015
UK	0.001	-0.001	0.007	-0.007	0.013	-0.013
Switzerland	0.001	-0.001	0.006	-0.006	0.012	-0.011
Netherlands	0.001	-0.001	0.006	-0.006	0.011	-0.011
USA	0.001	-0.001	0.005	-0.005	0.010	-0.009

Note: The table reports the estimated average indirect effects of distance and market size for the import-source country j on firms' exports to any destination (Ψ_{D_j} and Ψ_{Y_j}, respectively). These elasticities are computed at the firm-country level by multiplying the elasticity of χ_{hk} with respect to either Y ($\rho^f_{Y_j}$) or D ($\rho^f_{D_j}$) by the elasticity of exports to χ_{hk} obtained as the estimated coefficients on $\ln \widehat{\chi}^f_t$ reported in columns 7-9 of Table 2 for the extensive, intensive and both margins, respectively. All the estimated indirect effects are statistically significant at 1%. Standard errors, which are not reported, have been obtained by bootstrapping (500 replications).

$$\rho^f_{Y_j} = \frac{d\ln\chi_{hk}}{d\ln Y_j} = \frac{\alpha_h}{\phi_h - 1} * \frac{M^f_j}{\sum_{n=1}^N M^f_n} * \frac{d\ln\left(\frac{M^f_j}{M^f_k}\right)}{d\ln Y_j}. \tag{14}$$

The first term in both equations is the TFP elasticity to imports and can be retrieved from the estimates of the production function, column 4 of Table 1. The second element, which is directly observable in our data, is the fraction of imports of firm f from country j over the total intermediate inputs used by a firm. The third term can be obtained by estimating the elasticity of the ratio of imports from j over domestic intermediates with respect to distance and GDP. According to our theoretical setting, the ratio of imports of intermediates from country j to domestic intermediates can be expressed by

$$\frac{M^f_j}{M^f_k} = \frac{\beta_{mj}Y_j}{\beta_{mk}Y_k}\left(\left(\frac{w_j}{w_k}\right)\tau_{mjk}\right)^{1-\phi}.$$

Given that the above expression is log-linear in distance and market size, we first estimate by OLS the following equation[32]

$$\ln\frac{M^f_{jt}}{M^f_t} = a_0 + a_1\ln Y_{jt} + a_2\ln D_j + d^f + d_i + \epsilon^f_{jt}. \tag{15}$$

where, in addition to the two gravity forces Y_{jt} and D_j, we add a set of dummies to control for firm fixed-effects, d^f, and for year-geographical areas fixed-effects, d_i. Given that there might be other channels which are not explicitly included in our model such that exports affect imports (e.g., via complementarity between exports and imports in sunk/fixed cost or in transportation costs), as a robustness check we also run a specification including $Export^f_{jt}/DomesticSales^f_t$ as an additional explanatory variable. Finally, to take into account the large proportion of zeros observed in the data, we estimate the elasticity of the ratio with respect to gravity forces also by using a conditional (firm) fixed-effects Poisson regression.

The estimates of the log-linear specification are reported in columns 1 and 2 of Table 6, with and without the export share variable. In columns 3-4 we show the results of the Poisson

[32] Again, as before, to simplify the notation, we omit the subscript k when referring to the domestic country.

Table 8: Average indirect effects of gravity forces on export margins: by sector

Sector	Extensive Margin Ψ_{Y_h} (1)	Ψ_{D_h} (2)	Intensive Margin Ψ_{Y_h} (3)	Ψ_{D_h} (4)	Both Margins Ψ_{Y_h} (5)	Ψ_{D_h} (6)
Food, Beverages and Tobacco	0.018	-0.017	0.126	-0.122	0.299	-0.290
Textiles and Apparel	0.026	-0.025	0.186	-0.180	0.442	-0.428
Hide and Leather	0.028	-0.027	0.200	-0.193	0.475	-0.460
Wood and Cork	0.032	-0.031	0.226	-0.219	0.538	-0.521
Pulp and Paper	0.033	-0.032	0.234	-0.226	0.556	-0.539
Printing and Publishing	0.022	-0.021	0.154	-0.149	0.366	-0.354
Coke and Chemical products	0.032	-0.031	0.227	-0.219	0.539	-0.522
Rubber and Plastics	0.036	-0.035	0.256	-0.248	0.608	-0.589
Processing of non-metallic minerals	0.018	-0.017	0.127	-0.123	0.301	-0.291
Basic Metals	0.037	-0.035	0.260	-0.252	0.618	-0.599
Fabricated Metal Products	0.024	-0.023	0.168	-0.163	0.400	-0.387
Machinery and Equipment	0.010	-0.010	0.072	-0.070	0.171	-0.166
Electrical and Optical Equipment	0.017	-0.017	0.122	-0.118	0.291	-0.282
Motor Vehicles and Trailers	0.013	-0.013	0.094	-0.091	0.224	-0.217
Other Transport Equipment	0.031	-0.030	0.222	-0.215	0.528	-0.511
Other manufacturing industries	0.009	-0.009	0.067	-0.065	0.159	-0.154
All Manufacturing	0.018	-0.017	0.164	-0.159	0.315	-0.305

Note: The table reports the estimated average indirect effects for sector h of distance and market size on firms' exports to any destinationfor (Ψ_{D_h} and Ψ_{Y_h}, respectively). These elasticities are computed at the firm-level by multiplying the elasticity of χ_{hk} with respect to either Y (ρ^f_Y) or D (ρ^f_D) by the elasticity of exports to χ_{hk} obtained as the estimated coefficients on $\ln \widehat{\chi}^f_t$ reported in columns 7-9 of Table 2 for the extensive, intensive and both margins, respectively. All the estimated indirect effects are statistically significant at 1%. Standard errors, which are not reported, have been obtained by bootstrapping (500 replications).

regression, and we include as an additional control the proxy for fixed costs *Market Costs*. We observe that the elasticity of the import ratio is slightly lower than unity for both GDP and distance. Therefore, it is confirmed that firms' sourcing behaviour is influenced by the same standard gravity forces which are also active on the export side.

With the three terms of equations (13) and (14), we can now compute the indirect effect on a firm's export behaviour of the two gravity forces at the firm-origin level.

In Table 7 we report the estimated average indirect effects of distance and market size for the origin country j. These indirect effects are labeled Ψ_{D_j} and Ψ_{Y_j}, respectively. The table reports the results for the ten countries with the highest estimated effects. The results for market size and distance are quantitatively very similar, mainly due to the fact that the estimated elasticities of the import ratio with respect to the two gravity forces are almost identical (see Table 6). If we concentrate on the last two columns where we consider together both margins, the results indicate that, for firms importing from Germany, a rise in German market size of 10 percent would imply an increase of 0.7 percent in exports to each destination country.[33] A similar effect is detected for a decrease in transportation costs.

Together with the indirect effect of the two gravity forces for each import-source country j, it is possible to assess the indirect effect of a change in transportation costs or market size common across all countries. In this case, the elasticity of χ_{hk} with respect to market size (or distance) is given by

$$\rho^f_Y = \sum_{j \neq Italy} \rho^f_{Y_j} \quad \text{or} \quad \rho^f_D = \sum_{j \neq Italy} \rho^f_{D_j}$$

[33] As indicated in section 2.2, a change in transportation costs or market size of importing country j has an indirect impact on a firm's export behavior not only to country j but also to each export destination s, with $s \neq j$.

obtained by substituting the second element of equations (13) and (14) with the fraction of a firm's imports from all countries over its total intermediate inputs. The results in Table 8 show the average indirect effects of this generalized change for firms belonging to sector h. Some heterogeneity is observed, with the Rubber and Plastic and Basic Metals industries having the largest indirect impacts.

For the average manufacturing firm, the size of the estimated indirect effects of the gravity forces is about one third of the estimated direct effects obtained in the gravity equations (compare the last row of Table 8 with the estimated coefficients of GDP and distance reported in Table 2). Therefore, the magnitude of these indirect effects suggests that the TFP channel through which gravity forces affects exports is not just a theoretical possibility, but also an economically relevant mechanism. Our results confirm the predictions of the model according to which variations in trade costs and in the economic size of trade partners may have substantial indirect consequences on exporters' performance.

5. Conclusions

The recent heterogeneous-firm models have brought to the gravity model a need to consider the effects that the two gravity forces, namely market size and distance, have on firms' export patterns. This paper unveils a new channel through which these two forces affects firms international trade activities through their indirect effects on imports. Our theoretical framework introduces intermediate inputs into a Chaney (2008) model of trade with firm heterogeneity and asymmetric countries. The model shows that, in addition to a direct effect, market size and distance exert an additional effect on exports through the heterogeneous efficiency gains induced by imports of intermediate inputs. Indeed, importing has a positive effect on a firm's productivity which depends on both the mass of imported intermediate inputs available, as well as on the price of each intermediate. An increase in foreign market size has a positive effect on exports directly but also indirectly through an efficiency increase induced by the imports of intermediate inputs. Similarly, a decline in transportation costs, and therefore a reduction in the cost of imported inputs, has an indirect effect on a firm's exports pattern due to the increase in its productivity which allows to offer its exports at lower prices and to increase its revenues in the exporting markets.

The propositions of the model are tested using a large and unique panel data set of Italian manufacturing firms over the 2000-2006 period. First, we structurally estimate the contribution of importing to TFP. Second, we estimate how this improvement in efficiency affects firm-country margins of exports, controlling for the potential endogeneity of our key covariate. Third, we show that firms' import behaviour is affected by market size and distance and we quantify the indirect effect, via importing, of these two gravity forces on a firm's exports. We find that the elasticity of exports to market size and distance is magnified when imports of intermediates are accounted for: the size of the estimated indirect effects of the gravity forces is about one third of the estimated direct effects.

Overall, our findings suggest that the productivity gains from importing are heterogeneous depending on the import-source countries. The firms' productivity component due to imports has in turn a positive impact on firms' ability to sell their products internationally. Important policy implications follow from our results. Given that firms' sourcing strategies shape their export behavior, policies directly aimed at restricting imports by increasing trade costs or negative shocks occurring in import-source countries, can indirectly harm the export performance of domestic firms. Moreover, such events would impact more the most productive domestic firms, which make intensive use of imported inputs.

References

Acharya, R. C. and W. Keller (2009): "Technology transfer through imports," *Canadian Journal of Economics*, 42, 1411–1448.

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimator," *Econometrica*, 83, 2411–2451.

Agrawal, A., C. Rosell, and T. S. Simcoe (2014): "Do Tax Credits Affect R&D Expenditures by Small Firms? Evidence from Canada," NBER Working Papers 20615, National Bureau of Economic Research, Inc.

Aichele, R., G. J. Felbermayr, and I. Heiland (2014): "Going Deep: The Trade and Welfare Effects of TTIP," CESifo Working Paper Series 5150, CESifo Group Munich.

Altomonte, C. and G. Bekes (2009): "Trade Complexity and Productivity," IEHAS Discussion Papers 0914, Institute of Economics, Centre for Economic and Regional Studies, Hungarian Academy of Sciences.

Amiti, M., O. Itskhoki, and J. Konings (2014): "Importers, Exporters, and Exchange Rate Disconnect," *American Economic Review*, 104, 1942–78.

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

Anderson, J. E. and E. van Wincoop (2004): "Trade Costs," *Journal of Economic Literature*, 42, 691–751.

Antras, P., T. C. Fort, and F. Tintelnot (2017): "The margins of global sourcing: Theory and evidence from us firms," *American Economic Review*, 107, 2514–64.

Aristei, D., D. Castellani, and C. Franco (2013): "Firms' exporting and importing activities: is there a two-way relationship?" *Review of World Economics (Weltwirtschaftliches Archiv)*, 149, 55–84.

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121–68.

Baldwin, R. and J. Harrigan (2011): "Zeros, Quality, and Space: Trade Theory and Trade Evidence," *American Economic Journal: Microeconomics*, 3, 60–88.

Bas, M. (2012): "Input-trade liberalization and firm export decisions: Evidence from Argentina," *Journal of Development Economics*, 97, 481–493.

Bas, M. and V. Strauss-Kahn (2014): "Does importing more inputs raise exports? Firm level evidence from France," *Review of World Economics*, 150, 241–475.

——— (2015): "Input-trade liberalization, export prices and quality upgrading," *Journal of International Economics*, 95, 250–262.

Bernard, A. B., M. Grazzi, and C. Tomasi (2015): "Intermediaries in International Trade: Products and Destinations," *The Review of Economics and Statistics*, 97, 916–920.

Bernard, A. B. and J. B. Jensen (2004): "Why Some Firms Export," *The Review of Economics and Statistics*, 86, 561–569.

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

——— (2016): "Global Firms," NBER Working Papers 22727, National Bureau of Economic Research, Inc.

Blundell, R., R. Griffith, and F. Windmeijer (2002): "Individual effects and dynamics in count data models," *Journal of Econometrics*, 108, 113–131.

Bond, S. and M. Söderbom (2005): "Adjustment costs and the identification of Cobb Douglas production functions," Tech. rep., IFS Working Papers, Institute for Fiscal Studies (IFS).

Brandt, L., J. Van Biesebroeck, L. Wang, and Y. Zhang (2012): "WTO Accession and Performance of Chinese Manufacturing Firms," CEPR Discussion Papers 9166, C.E.P.R. Discussion Papers.

Caliendo, L. and F. Parro (2015): "Estimates of the Trade and Welfare Effects of NAFTA," *Review of Economic Studies*, 82, 1–44.

Castellani, D., F. Serti, and C. Tomasi (2010): "Firms in International Trade: Importers' and Exporters' Heterogeneity in Italian Manufacturing Industry," *The World Economy*, 33, 424–457.

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

Chevassus-Lozza, E., C. Gaignè, and L. Le Mener (2013): "Does input trade liberalization boost downstream firms' exports? Theory and firm-level evidence," *Journal of International Economics*, 90, 391–402.

Conti, G., A. LoTurco, and D. Maggioni (2014): "Rethinking the import-productivity nexus for Italian manufacturing," *Empirica*, 41, 589–617.

Creusen, H., H. Kox, A. Lejour, and R. Smeets (2011): "Exploring the Margins of Dutch Exports: A Firm-Level Analysis," *De Economist*, 159, 413–434.

Crozet, M. and P. Koenig (2010): "Structural gravity equations with intensive and extensive margins," *Canadian Journal of Economics*, 43, 41–62.

Dauth, W., S. Findeisen, and J. Suedekum (2018): "Adjusting to Globalization in Germany," IZA Discussion Papers 11299, Institute for the Study of Labor (IZA).

De Loecker, J. (2013): "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, 5, 1–21.

De Sousa, J., T. Mayer, and S. Zignago (2012): "Market access in global and regional trade," *Regional Science and Urban Economics*, 42, 1037–1052.

Djankov, S., C. Freund, and C. S. Pham (2010): "Trading on time," *Review of Economics and Statistics*, 92, 166–173.

Donoso, V., V. Martin, and A. Minondo (2015): "Do Differences in the Exposure to Chinese Imports Lead to Differences in Local Labour Market Outcomes? An Analysis for Spanish Provinces," *Regional Studies*, 49, 1746–1764.

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

Eaton, J., S. S. Kortum, and F. Kramarz (2011): "An Anatomy of International Trade: Evidence from French Firms," *Econometrica*, 79, 1453–1498.

Fan, H., E. L.-C. Lai, and H. S. Qi (2016): "Imported-input Trade Liberalization and Firms' Export Performance in China: Theory and Evidence," CESifo Working Paper Series 5828, CESifo Group Munich.

Feng, L., Z. Li, and D. L. Swenson (2016): "The Connection between Imported Intermediate Inputs and Exports: Evidence from Chinese Firms," *Journal of International Economics*, 101, 86–101.

Fernandes, A. M. (2007): "Trade policy, trade volumes and plant-level productivity in Colombian manufacturing industries," *Journal of International Economics*, 71, 52–71.

Gandhi, A., S. Navarro, and D. Rivers (2018): "On the Identification of Gross Output Production Functions," University of Western Ontario, Centre for Human Capital and Productivity (CHCP) Working Papers 20181, University of Western Ontario, Centre for Human Capital and Productivity (CHCP).

Gandhi, A., S. Navarro, D. Rivers, et al. (2017): "How heterogeneous is productivity? A comparison of gross output and value added," *Journal of Political Economy, forthcoming.*

Gorg, H., A. Hanley, and E. Strobl (2008): "Productivity effects of international outsourcing: evidence from plant-level data," *Canadian Journal of Economics*, 41, 670–688.

Grazzi, M., R. Sanzo, A. Secchi, and A. Zeli (2013): "The building process of a new integrated system of business micro-data 1989-2004," *Journal of Economic and Social Measurement*, 291–324.

Grossman, G. M. and E. Helpman (1991): *Innovation and Growth in the Global Economy*, Cambridge, Mass. ; London : MIT.

Gwartney, J., R. Lawson, and J. Hall (2014): "2014 Economic Freedom Dataset, published in Economic Freedom of the World: 2014 Annual Report," *Fraser Institute.*

Halpern, L., M. Koren, and A. Szeidl (2015): "Imported inputs and productivity," *American Economic Review*, 105, 3660–3703.

Head, K. and T. Mayer (2014): *Gravity Equations: Workhorse,Toolkit, and Cookbook*, Elsevier, vol. 4 of *Handbook of International Economics*, chap. 0, 131–195.

Helpman, E., M. Melitz, and Y. Rubinstein (2008): "Estimating Trade Flows: Trading Partners and Trading Volumes," *The Quarterly Journal of Economics*, 123, 441–487.

Hummels, D., R. J?rgensen, J. Munch, and C. Xiang (2014): "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data," *American Economic Review*, 104, 1597–1629.

Kasahara, H. and B. Lapham (2013): "Productivity and the decision to import and export: Theory and evidence," *Journal of International Economics*, 89, 297–316.

Kasahara, H. and J. Rodrigue (2008): "Does the use of imported intermediates increase productivity? Plant-level evidence," *Journal of Development Economics*, 87, 106–118.

Krugman, P. (1980): "Scale economies, product differentiation, and the pattern of trade," *The American Economic Review*, 70, 950–959.

Lawless, M. and K. Whelan (2014): "Where Do Firms Export, How Much and Why?" *The World Economy*, 37, 1027–1050.

Levinsohn, J. and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–41.

Loof, H. and M. Andersson (2010): "Imports, Productivity and Origin Markets: The Role of Knowledge-intensive Economies," *The World Economy*, 33, 458–481.

LoTurco, A. and D. Maggioni (2013): "On the Role of Imports in Enhancing Manufacturing Exports," *The World Economy*, 36, 93–120.

——— (2015): "Imports, Exports and the Firm Product Scope: Evidence From Turkey," *The World Economy*, 38, 984–1005.

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

Markusen, J. (1989): "Trade in Producer Services and in Other Specialized Intermediate Inputs," *American Economic Review*, 79, 85–95.

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

Mion, G. and L. Zhu (2013): "Import competition from and offshoring to China: A curse or blessing for firms?" *Journal of International Economics*, 89, 202–215.

Muuls, M. and M. Pisu (2009): "Imports and Exports at the Level of the Firm: Evidence from Belgium," *The World Economy*, 32, 692–734.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Paul, C. J. M. and M. Yasar (2009): "Outsourcing, productivity, and input composition at the plant level," *Canadian Journal of Economics*, 42, 422–439.

Pavcnik, N. (2002): "Trade Liberalization, Exit, and Productivity Improvement: Evidence from Chilean Plants," *Review of Economic Studies*, 69, 245–76.

Rivera-Batiz, L. A. and P. M. Romer (1991): "International trade with endogenous technological change," *European Economic Review*, 35, 971–1001.

Romer, P. M. (1990): "Endogenous Technological Change," *Journal of Political Economy*, 98, S71–102.

Serti, F. and C. Tomasi (2014): "Export and import market-specific characteristics," *Empirical Economics*, 47, 1467–1496.

Silva, J. M. C. S. and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658.

Tinbergen, J. (1962): *Shaping the World Economy: sugggestions for an International Economic Policy*, Twentieth Century Fund, New-York.

Vogel, A. and J. Wagner (2010): "Higher productivity in importing German manufacturing firms: self-selection, learning from importing, or both?" *Review of World Economics (Weltwirtschaftliches Archiv)*, 145, 641–665.

Yi, K.-M. (2003): "Can Vertical Specialization Explain the Growth of World Trade?" *Journal of Political Economy*, 111, 52–102.

Technical Appendix.
The role of the gravity forces on firms' trade

Antonio Navas[a], Francesco Serti[b], Chiara Tomasi[c,*]

[a]*Department of Economics, University of Sheffield*
[b]*FAE, Universidad de Alicante and AXES IMT, Lucca*
[c]*Department of Economics and Management, University of Trento and LEM, Pisa*

Abstract

This paper offers both a theoretical framework and empirical evidence on the role that the two gravity forces, namely market size and geographical distance, have indirectly through imports, on firms' exports patterns. The model shows that sourcing from bigger and closer markets implies higher productivity gains which, in turn, increase firms' ability to enter export market, as well as their export value. Exploiting data on product and destination-level transactions of a large panel of Italian firms, the paper shows that, on average, the indirect effects of the gravity forces are about one third of their direct effects.

JEL codes: F12, F14.
Keywords: Gravity Forces, Geographical distance, Market size, Firms' heterogeneity, Imports and Exports.

*Corresponding author. *Email:* chiara.tomasi@unitn.it, *Address*: Department of Economics and Management, Via Inama 5, 38122, Trento, Italy. *Tel* +39 (0)461 282161

Email addresses: a.navas@sheffield.ac.uk (Antonio Navas), serti@gmail.com (Francesco Serti)

1. Introduction

This technical appendix contains additional theoretical and empirical material that complement the results shown in the paper "The role of the gravity forces on firms' trade". In Section 2, we present a detailed analysis of the solution of our model which includes the full set of derivations and extensions that reinforce the results of the paper. In Section 3 we provide additional data information and empirical analyses.

2. Additional theoretical parts

2.1. Solution of the model

In the paper we have described the main elements and assumptions of our theoretical model and we have disclosed the main theoretical results. In this section we explain in detail how those results have been derived.

First, we start by studying how firms choose the optimal combination of inputs for a given production target. As standard in micro-economic theory, this can be done by focusing on the firms' cost minimization. In this framework, this problem can be solved in two steps. In the first step, the firm can select the optimal combination of intermediate input varieties for a given firm's demand of the intermediate composite good m^f_{hk}. Then in a second step the firm can choose the optimal combination of labor and the intermediate composite good, for a given production quantity q^f_{hk}.

Focusing on interior solutions, the first step involves solving the following problem,

$$Min \int\limits_{\nu \epsilon \Lambda} p_{hmk}(\nu) m^f_{hk}(\nu)\, d\nu$$

$$s.t. \left(\int\limits_{\nu \epsilon \Lambda} \left(m^f_{hk}(\nu) \right)^{\frac{\phi_h - 1}{\phi_h}} d\nu \right)^{\frac{\phi_h}{\phi_h - 1}} = m^f_{hk} \quad .$$

This leads to the standard demand function for each intermediate input

$$m^f_{hk}(\nu) = m^f_{hk} \left(\frac{p_{hmk}(\nu)}{P_{hmk}} \right)^{-\phi_h}$$

where the aggregate price index for the intermediate composite good is given by

$$P_{hmk} = \left(\int\limits_{\nu \epsilon \Lambda} \left(p_{hmk}(\nu) \right)^{1-\phi_h} d\nu \right)^{\frac{1}{1-\phi_h}} \quad .$$

We assume that the mass of intermediate input varieties available is different across countries. Since each intermediate producer is a monopolist, then each firm will charge $p_{hmk}(\nu) = \frac{\phi_h \tau_{mjk} w_j}{\phi_h - 1}$ where $\tau_{mkk} = 1$. Applying symmetry across all intermediate inputs belonging to the same country, we can express the aggregate price index for the intermediate composite good in country k and sector h as

$$P_{hmk} = \left(\sum_{j=1}^{N} \left(w_j \tau_{mjk} \right)^{1-\phi_h} \tilde{L}_j \right)^{\frac{1}{1-\phi_h}} \frac{\phi_h}{\phi_h - 1}$$

where $\tilde{L}_j = \beta_{mj} w_j L_j$, represents the mass of intermediate input varieties coming from country j.

In the second step, a firm chooses the optimal combination of labor and the intermediate composite good by solving the following problem

$$Min\ w_k l^f_{hk} + P_{hmk} m^f_{hk}$$

$$s.t. q^f_{hk} = \varphi^f \left(m^f_{hk}\right)^{\alpha_h} \left(l^f_{hk}\right)^{1-\alpha_h}$$

The conditional demand for each input of a firm with productivity φ^f is given by

$$l^f_{hk} = \frac{1}{\varphi^f}\left(\frac{P_{hmk}}{w_k}\frac{1-\alpha_h}{\alpha_h}\right)^{\alpha_h} q^f_{hk}$$

$$m^f_{hk} = \frac{1}{\varphi^f}\left(\frac{w_k}{P_{hmk}}\frac{\alpha_h}{1-\alpha_h}\right)^{1-\alpha_h} q^f_{hk} \quad .$$

Substituting the last two equations in the objective function we obtain the variable cost function for a firm with innate productivity φ^f in country k and sector h

$$c^f_{hk}\left(\varphi^f\right) = \frac{(w_k)^{1-\alpha_h}(P_{hmk})^{\alpha_h}}{\Gamma_h}\frac{q^f_{hk}}{\varphi^f} = \frac{(\rho_{hm})^{\alpha_h} w_k}{\Gamma_h (\chi_{hk})^d \left(\tilde{L}_k\right)^{\frac{\alpha_h}{\phi_h-1}}}\frac{q^f_{hk}}{\varphi^f}$$

where $d = 1$ if a firm imports intermediates (and 0 otherwise) and $\Gamma_h = \alpha_h^{\alpha_h}(1-\alpha_h)^{1-\alpha_h}$. We denote with $\rho_{hm} = \frac{\phi_h}{\phi_h - 1}$ the mark-up of the intermediate producers.

The variable c^f_{hk} is a linear function of

$$\chi_{hk} = \left[\sum_{j=1}^{N}\left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h}\frac{\tilde{L}_j}{\tilde{L}_k}\right]^{\frac{\alpha_h}{\phi_h-1}}$$

which, as explained in the main text and below, captures the contribution of importing intermediate inputs to a firms' total factor productivity (TFP) and it is a key element in our analysis. The variable Γ_h is a technological constant and $\tilde{L}_k = \beta_{mk} w_k L_k$. Notice that for an importing firm $\chi_{hk} > 1$. It follows that an importer enjoys lower marginal costs of production. Note that, not all sourcing countries contribute equally to the increase in productivity that follows from importing. This is crucial in our analysis as we will comment below.

Once obtained the cost function of the firm, we proceed by solving the firms' profit maximization problem in the final good sector. As usual in the Dixit Stiglitz monopolistic competition framework, the price set by a firm is a constant mark-up over the marginal cost of production. Therefore, the price on market j of a final good produced in country k by a firm with productivity φ^f is

$$p^f_{hxkj}\left(\varphi^f\right) = \frac{\sigma_h}{\sigma_h - 1}\frac{(\rho_{hm})^{\alpha_h}}{\Gamma_h \chi_{hk}\left(\tilde{L}_k\right)^{\frac{\alpha_h}{\phi_h-1}}}\frac{\tau_{xkj} w_k}{\varphi^f} \quad .$$

Substituting the price expression in the demand function we obtain the quantity sold in country j by a final good producer of country k, which is

$$q^f_{hxkj}\left(\varphi^f\right) = \frac{\mu_h R_j}{(P_{hj})^{1-\sigma_h}}\left(\frac{\tau_{xkj}\rho_h (\rho_{hm})^{\alpha_h} w_k}{\Gamma_h \chi_{hk}\left(\tilde{L}_k\right)^{\frac{\alpha_h}{\phi_h-1}}\varphi^f}\right)^{-\sigma_h},$$

where $\rho_h = \frac{\sigma_h}{\sigma_h - 1}$ is the mark-up of final goods producers belonging to sector h. For a firm belonging to sector h of country k, the operating profits from selling to country j are given by

$$r^f_{hxkj}\left(\varphi^f\right) = \left(\tau_{xkj}\right)^{1-\sigma_h} \frac{\mu_h R_j}{\sigma_h \left(P_{hj}\right)^{1-\sigma_h}} \left(\frac{\rho_h \left(\rho_{hm}\right)^{\alpha_h} w_k}{\Gamma_h \chi_{hk} \left(\tilde{L}_k\right)^{\frac{\alpha_h}{\phi_h - 1}} \varphi^f} \right)^{1-\sigma_h} .$$

A firm of country k will export to country j when the operating profits of serving country j overcome the fixed cost of exporting (i.e. $r_{hxkj}(\varphi^f) \geq F_{hxkj}$). Since the export operating profits are monotonically increasing in the firms' innate productivity, there exists a productivity threshold φ^*_{hxkj} such that if the firms' innate productivity is above this threshold, the firm will export to country j. The expression for this productivity cut-off is given by the equation

$$\varphi^*_{hxkj} = \tau_{xkj} \left(\frac{\sigma_h F_{hxkj}}{\mu_h R_j} \right)^{\frac{1}{\sigma_h - 1}} \frac{\rho_h w_k}{P_{hj}} \frac{\left(\rho_{hm}\right)^{\alpha_h} \left(\tilde{L}_k\right)^{\frac{\alpha_h}{1-\phi_h}}}{\chi_{hk} \Gamma_h} . \tag{1}$$

A firm of country k will import when the gain in operating profits from importing overcome the fixed costs of importing. As explained in the main manuscript, we focus on an equilibrium in which the firm indifferent between importing or relying on domestic intermediate inputs is a domestic non-exporting firm.[1] The operating profits of a domestic firm are given by

$$r^f_{hk}(\varphi^f) = \frac{\mu_h R_k}{\sigma_h \left(P_{hk}\right)^{1-\sigma_h}} \left(\frac{\psi_h w_k}{\chi^d_{hk} \left(\tilde{L}_k\right)^{\frac{\alpha_h}{\phi_h - 1}} \varphi^f} \right)^{1-\sigma_h} .$$

where $\psi_h = \frac{\rho_h (\rho_{hm})^{\alpha_h}}{\Gamma_h}$. Note that $r^f_{hik}(\varphi^f) = \left(\chi_{hk}\right)^{\sigma_h - 1} r^f_{hk}\left(\varphi^f\right)$, where $r^f_{hik}(\varphi^f)$, represents the operating profits of a domestic firm that imports. Therefore, a firm in k will be importing intermediates if $r^f_{hik}(\varphi^f) - r^f_{hk}(\varphi^f) \geq F_{hik.}$ The productivity threshold associated to importing is given by the following expression

$$\varphi^*_{hik} = \left(\left(\chi_{hk}\right)^{\sigma_h - 1} - 1 \right)^{\frac{1}{1-\sigma_h}} \left(\frac{\sigma_h F_{hik}}{\mu_h R_k} \right)^{\frac{1}{\sigma_h - 1}} \frac{\psi_h w_k}{P_{hk}} \left(\tilde{L}_k\right)^{\frac{\alpha_h}{1-\phi_h}} . \tag{2}$$

This expression indicates that the larger the gains from importing, (i.e. the larger χ_{hk}), the lower the import productivity threshold. Moreover, the larger the home market, R_k, the lower the productivity threshold and, therefore, the larger the mass of importing firms.[2]

Finally, the survival productivity threshold is described by the following equation

$$\varphi^*_{hk} = \left(\frac{\sigma_h F_h}{\mu_h R_k} \right)^{\frac{1}{\sigma_h - 1}} \frac{\psi_h w_k}{P_{hk}} \left(\tilde{L}_k\right)^{\frac{\alpha_h}{1-\phi_h}} . \tag{3}$$

Note that these productivity thresholds depend on the aggregate price index, P_{hj} which is

[1] A sufficient and necessary condition for the existence of this equilibria is $\left(\tau_{xkj}\right)^{\sigma_h - 1} \left(\frac{Y_k}{Y_j}\right)^{\frac{\sigma_h - 1}{\gamma_h}} \left(\frac{\theta'_{hk}}{\theta'_{hj}}\right)^{\sigma_h - 1} \frac{(\chi_{hk})^{\sigma_h - 1} - 1}{(\chi_{hk})^{\sigma_h - 1}} F_{hxkj} \geq F_{hik} \geq \left(\left(\chi_{hk}\right)^{\sigma_h - 1} - 1 \right) F_h \quad \forall \quad j.$

[2] This is due to two different mechanisms. First, a larger home market, R_k, implies a larger demand of final goods and, as a consequence, a larger demand of intermediate inputs. Second, firms in larger markets have access to a larger set of intermediate inputs and, therefore, have a lower marginal cost. As the gains from importing intermediates are inversely proportional to the marginal cost of production, firms' profits from importing intermediates are larger in larger markets.

an endogenous variable. The expression for the aggregate price index is given by

$$P_{hj}^{1-\sigma_h} = \underbrace{\beta_{hj} w_j L_j \int_{\varphi_{hj}^*}^{\infty} (p_{hj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi}_{\text{Domestic firms}} + \underbrace{\sum_{n \neq j}^{N} \beta_{hn} w_n L_n \int_{\varphi_{hxnj}^*}^{\infty} (p_{hxnj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi}_{\text{foreign exporters}} \quad .$$

To be able to derive the gravity equation, an expression for the aggregate price index is needed. In contrast to models in which firms are not allowed to import, we need to distinguish between domestic importers and non-importers, as they price differently

$$\int_{\varphi_{hj}^*}^{\infty} (p_{hj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi = \int_{\varphi_{hj}^*}^{\varphi_{hij}^*} (p_{hj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi + \int_{\varphi_{hij}^*}^{\infty} (p_{hij}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi.$$

In the following steps we compute each of these integrals. Substituting the expressions for $p_{hj}(\varphi), p_{hij}(\varphi)$ we have that

$$\int_{\varphi_{hj}^*}^{\infty} (p_{hj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi = \left(w_j \left(\tilde{L}_j\right)^{\frac{\alpha_h}{1-\phi_h}} \psi_j \right)^{1-\sigma_h}$$
$$\times \underbrace{\left(\int_{\varphi_{hj}^*}^{\varphi_{hij}^*} \varphi^{\sigma_h - 1} g(\varphi)\, d\varphi + (\chi_{hk})^{\sigma_h - 1} \int_{\varphi_{hij}^*}^{\infty} \varphi^{\sigma_h - 1} g(\varphi)\, d\varphi. \right)}_{A} \quad .$$

Taking the derivative of the cumulative distribution function we obtain the density function $g(\varphi) = \gamma_h (\varphi)^{-(\gamma_h + 1)}$. Substituting in the latter expression and solving for the integrals we have that

$$A = \frac{\gamma_h}{\gamma_h - (\sigma_h - 1)} \left[\left(\varphi_{hj}^*\right)^{\sigma_h - \gamma_h - 1} + \left(\varphi_{hij}^*\right)^{\sigma_h - \gamma_h - 1} \left((\chi_{hk})^{\sigma_h - 1} - 1 \right) \right] \quad .$$

Dividing 2 by 3 we find that $\left(\frac{\varphi_{hij}^*}{\varphi_{hj}^*}\right) = \left(\frac{F_{hik}}{\left((\chi_{hk})^{\sigma_h - 1} - 1\right) F_h} \right)^{\frac{1}{h-1}}$. Rearranging terms yields

$$\int_{\varphi_{hj}^*}^{\infty} (p_{hj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi = \left(w_j \left(\tilde{L}_j\right)^{\frac{\alpha_h}{1-\phi_h}} \psi_j \right)^{1-\sigma_h} \frac{\gamma_h}{\gamma_h - (\sigma_h - 1)}$$
$$\times \left[\frac{(F_h)^{\frac{\sigma_h - \gamma_h - 1}{\sigma_h - 1}} + \left((\chi_{hk})^{\sigma_h - 1} - 1\right)^{\frac{\gamma_h}{\sigma_h - 1}} (F_{hik})^{\frac{\sigma_h - \gamma_h - 1}{\sigma_h - 1}}}{(F_h)^{\frac{\sigma_h - \gamma_h - 1}{\sigma_h - 1}}} \right] \left(\varphi_{hj}^*\right)^{\sigma_h - \gamma_h - 1} .$$

Substituting the expression for φ_{hj}^* obtained from equation (3), and rearranging terms

$$\int_{\varphi_{hj}^*}^{\infty} (p_{hj}(\varphi))^{1-\sigma_h} g(\varphi)\, d\varphi = \left(w_j \left(\tilde{L}_j\right)^{\frac{\alpha_h}{1-\phi_h}} \psi_j \right)^{1-\sigma_h} \frac{\gamma_h}{\gamma_h - (\sigma_h - 1)}$$
$$\times \left[(F_h)^{\frac{\sigma_h - \gamma_h - 1}{\sigma_h - 1}} + \left((\chi_{hk})^{\sigma_h - 1} - 1\right)^{\frac{\gamma_h}{\sigma_h - 1}} (F_{hik})^{\frac{\sigma_h - \gamma_h - 1}{\sigma_h - 1}} \right]$$
$$\times \left(\frac{\sigma_h}{\mu_h R_k} \right)^{\frac{\sigma_h - \gamma_h - 1}{\sigma_h - 1}} \left(\frac{\psi_h w_j}{P_{hj}} \right)^{\sigma_h - \gamma_h - 1} \left(\tilde{L}_k\right)^{\frac{\alpha_h(\sigma_h - \gamma_h - 1)}{1-\phi_h}} \quad .$$

Now we compute the foreign exporters part. Substituting for the optimal prices and rearranging terms we have that

$$\sum_{n\neq j}^{N}\beta_{hn}w_nL_n\int\limits_{\varphi^*_{hxnj}}^{\infty}(p_{hxnj}(\varphi))^{1-\sigma_h}g(\varphi)\,d\varphi=\sum_{n\neq j}^{N}\beta_{hn}w_nL_n\left(\frac{\psi_h\tau_{xnj}w_n}{\chi_{hn}\left(\tilde{L}_n\right)^{\frac{\alpha_h}{\phi_h-1}}}\right)^{1-\sigma_h}\times\int\limits_{\varphi^*_{hxnj}}^{\infty}(\varphi)^{\sigma_h-1}g(\varphi)\,d\varphi\quad.$$

Solving for the integral we have that

$$\sum_{n\neq j}^{N}\beta_{hn}w_nL_n\left(\frac{\psi_h\tau_{xnj}w_n}{\chi_{hn}\left(\tilde{L}_n\right)^{\frac{\alpha_h}{\phi_h-1}}}\right)^{1-\sigma_h}\left(\frac{\gamma_h}{\gamma_h-(\sigma_h-1)}\right)\left(\varphi^*_{hxnj}\right)^{\sigma_h-\gamma_h-1}$$

and substituting the expression for the exporting productivity cutoff and rearranging terms yields

$$\sum_{n\neq j}^{N}\beta_{hn}w_nL_n\left(\frac{\psi_h\tau_{xnj}w_n}{\chi_{hn}\left(\tilde{L}_n\right)^{\frac{\alpha_h}{\phi_h-1}}}\right)^{1-\sigma_h}\left(\frac{\gamma_h}{\gamma_h-(\sigma_h-1)}\right)\left(\frac{\sigma_hF_{hxnj}}{\mu_hR_j}\right)^{\frac{\sigma_h-\gamma_h-1}{\sigma_h-1}}\times\left(\frac{\tau_{xnj}\psi_hw_n}{P_{hj}\chi_{hn}\left(\tilde{L}_n\right)^{\frac{\alpha_h}{\phi_h-1}}}\right)^{\sigma_h-\gamma_h-1}$$

Putting both integrals together, and rearranging terms

$$\begin{aligned}P_{hj}^{-\gamma_h}=&\left(\frac{\gamma_h}{\gamma_h-(\sigma_h-1)}\right)\left(\frac{\sigma_h}{\mu_hR_j}\right)^{\frac{\sigma_h-\gamma_h-1}{\sigma_h-1}}\\&\times\sum_{n=1}^{N}\beta_{hn}w_nL_n\left(\tilde{L}_n\right)^{\frac{\alpha_h\gamma_h}{\phi_h-1}}(\psi_hw_n\tau_{xnj})^{-\gamma_h}\left(\chi_{hn}^{\gamma_h}(F_{hxnj})^{\left(\frac{\sigma_h-\gamma_h-1}{\sigma_h}\right)}\right)^{(1-\xi)}(\Phi_h)^{\xi}\end{aligned}$$

where $\Phi_h=(F_h)^{\left(\frac{\sigma_h-\gamma_h-1}{\sigma_h-1}\right)}+\left((\chi_{hn})^{\sigma_h-1}-1\right)^{\frac{\gamma_h}{\sigma_h-1}}(F_{hin})^{\left(\frac{\sigma_h-\gamma_h-1}{\sigma_h-1}\right)}$ and ξ is an indicator function taking the value of 1 if $n=j$ and 0 otherwise. Defining $\lambda_{2h}^{'\gamma_h}=\left(\frac{\gamma_h-(\sigma_h-1)}{\gamma_h}\right)\left(\frac{\sigma_h}{\mu_h}\right)^{\frac{\sigma_h-\gamma_h-1}{1-\sigma_h}}\left(\frac{1+\pi}{Y}\right)$ and taking into account that $R_j=w_jL_j(1+\pi)=Y_j$, and rearranging terms, P_{hj} can be expressed as

$$\begin{aligned}P_{hj}&=\lambda'_{2h}(Y_j)^{\frac{1}{\gamma_h}-\frac{1}{\sigma_h-1}}\theta'_{hj}\\\left(\theta'_{hj}\right)^{-\gamma_h}&=\left[\sum_{n=1}^{N}\frac{Y_n}{Y}(w_n\tau_{xnj}\psi_h)^{-\gamma_h}\left(\chi_{hn}^{\gamma_h}(F_{hxnj})^{\frac{\sigma_h-\gamma_h-1}{\sigma_h}}\right)^{1-\xi}\beta_{hn}\left(\tilde{L}_n\right)^{\frac{\alpha_h\gamma_h}{\phi_h-1}}(\Phi_h)^{\xi}\right]\quad.\end{aligned}$$

where θ'_{hj} is the multilateral resistance term, which takes also into account the fact that some firms are importing intermediate inputs and, consequently, they are charging different prices; λ'_{2h} is a constant term. In what follows we assume that our country is a small open economy. This implies that any change in the domestic market does not have any relevant impact on the measure θ'_{hj}. This simplifies significantly the calculations. With the definition of the price index in hand, we are able to derive the general equilibrium value of the export productivity cutoffs and of firm-level exports.

Plugging P_{hj} in (1) and using again the fact that $R_j = Y_j$, we obtain the equilibrium value of the productivity threshold for exports. With that we obtain the two gravity equations as shown in the paper,

$$\Pr(\varphi \geq \varphi^*_{hxkj}) = (\lambda'_{4h})^{-\gamma_h} \underbrace{\left(\frac{Y_j}{Y}\right)\left(\frac{w_k \tau_{xkj}}{\theta'_{hj}}\right)^{-\gamma_h} (F_{hxkj})^{\frac{-\gamma_h}{\sigma_h-1}}}_{Chaney's} \underbrace{(\tilde{\chi}_{hk})^{\gamma_h}}_{intermediate\ contribution} \tag{4}$$

$$\begin{aligned} X^f_{hxkj}(\varphi^f) &= \left(\frac{\varphi^f}{\varphi^*_{hxkj}}\right)^{\sigma_h-1} \sigma_h r_{hxkj}(\varphi^*_{hxkj}) \\ &= \lambda'_{3h} \underbrace{\left(\frac{Y_j}{Y}\right)^{\frac{\sigma_h-1}{\gamma_h}} \left(\frac{\theta'_{hj}}{w_k \tau_{xkj}}\right)^{\sigma_h-1}}_{Chaney's} \underbrace{(\tilde{\chi}_{hk})^{\sigma_h-1}}_{intermediate\ contribution} (\varphi^f)^{\sigma_h-1} . \end{aligned}$$

where λ'_{3h} and λ'_{4h} are constants.[3]

2.2. Imported intermediate inputs and firms' productivity

Proposition 1 in the paper states that importing intermediate inputs has a positive effect on a firm's productivity. This effect depends on the characteristics of the country of origin of imports. To formalize the intuition behind these results, we can derive a firm's total factor productivity (TFP). The demand of a firm in country k for an intermediate input produced in country j can be expressed as

$$m^f_{hj}(\nu) = \left(\frac{p_{hmj}(\nu)}{p_{hmk}(\nu)}\right)^{-\phi_h} m^f_{hk}(\nu) \tag{5}$$

where $m^f_{hk}(\nu)$ is the demand for a domestic variety. Note that there is symmetry across all varieties from a specific location. Therefore in equilibrium $m^f_{hj}(\nu) = m^f_{hj}(\nu')\ \forall\ \nu, \nu' \in \Lambda$). Let us denote with $\bar{m}^f_{hk}$ the quantity used of each domestic intermediate input variety in equilibrium and p_{hmk} its associated price.

The total volume of intermediate inputs used by a firm, M^f_{tot}, can be expressed as

$$M^f_{tot} = \int_{\nu\epsilon\Lambda} \frac{p_{hmj}(\nu)\, m^f_{hj}(\nu)}{p_{hmk}} d\nu = \left[\sum_{j=1}^{N} \left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h} \frac{\tilde{L}_j}{\tilde{L}_k}\right] \left(\tilde{L}_k\right) \bar{m}^f_{hk} \tag{6}$$

where the expression above is obtained by using 5 and the symmetry condition above. Notice that M^f_{tot} is the value of the intermediate inputs used by a firm deflated by the domestic intermediate input price.

From the main manuscript, we obtain that the definition for the intermediate composite good is given by

$$m^f_{hk} = \left(\int_{v\epsilon\Lambda} \left(m^f_{hj}(\nu)\right)^{\frac{\phi_h-1}{\phi_h}} d\nu\right)^{\frac{\phi_h}{\phi_h-1}} \tag{7}$$

Substituting equation 5 and the expression for the optimal prices for each intermediate input

[3] $\lambda'_{4h} = \left(\frac{\gamma_h}{\gamma_h-(\sigma_h-1)}\right)^{\frac{1}{\gamma_h}} \left(\frac{\sigma_h}{\mu_h}\right)^{\frac{1}{\gamma_h}} (1+\pi)^{\frac{-1}{\gamma_h}} \psi_h \left(\frac{1+\pi}{Y}\right)^{\frac{\alpha_h}{1-\phi_h}}$ and $\lambda'_{3h} = \sigma_h (\lambda'_{4h})^{1-\sigma_h}$.

source and rearranging terms we obtain that

$$m^f_{hk} = \left[\sum_{j=1}^{N}\left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h}\tilde{L}_j\right]^{\frac{\phi_h}{\phi_h-1}} \bar{m}_{hk} \tag{8}$$

Rearranging terms we obtain

$$m^f_{hk} = \left[\sum_{j=1}^{N}\left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h}\frac{\tilde{L}_j}{\tilde{L}_k}\right]^{\frac{\phi_h}{\phi_h-1}} \tilde{L}_k^{\frac{\phi_h}{\phi_h-1}}\bar{m}_{hk} \tag{9}$$

Therefore, 9 can be expressed as the total volume of imports multiplied by a weighted average of country characteristics

$$m^f_{hk} = M^f_{tot}\left[\sum_{j=1}^{N}\left(\left(\frac{w_j}{w_k}\right)\tau_{mjk}\right)^{1-\phi_h}\frac{\tilde{L}_j}{\tilde{L}_k}\right]^{\frac{1}{\phi_h-1}} \left(\tilde{L}_k\right)^{\frac{1}{\phi_h-1}}$$

and by plugging this equation into the production function

$$q^f_{hn} = \varphi^f_h\left(l^f_{hn}\right)^{1-\alpha_h}\left(m^f_{hn}\right)^{\alpha_h}$$

we get

$$q^f_{hk} = \varphi^f_h\left(l^f_{hk}\right)^{1-\alpha_h}\left(M^f_{tot}\right)^{\alpha_h}\left[\sum_{j=1}^{N}\left(\left(\frac{w_j}{w_k}\right)\tau_{mjk}\right)^{1-\phi_h}\frac{\tilde{L}_j}{\tilde{L}_k}\right]^{\frac{\alpha_h}{\phi_h-1}}\left(\tilde{L}_k\right)^{\frac{\alpha_h}{\phi_h-1}} \quad .$$

The last two terms reflect the gains from variety obtained from imported and domestic intermediate inputs, respectively. By expressing the mass of each country intermediate input varieties as a function of GDP,

$$\tilde{L}_j = \beta_{mj}w_jL_j = \beta_{mj}\frac{Y_j}{(1+\pi)}$$

and rearranging terms we obtain,

$$\frac{q^f_{hk}}{\left(l^f_{hk}\right)^{1-\alpha_h}\left(M^f_{tot}\right)^{\alpha_h}} = \varphi^f_h\underbrace{\underbrace{\left[\sum_{j=1}^{N}\left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h}\frac{\beta_{mj}}{\beta_{mk}}\frac{Y_j}{Y_k}\right]^{\frac{\alpha_h}{\phi_h-1}}}_{\chi_{hk}}\left(\frac{\beta_{mk}Y_k}{Y}\right)^{\frac{\alpha_h}{\phi_h-1}}\left(\frac{(1+\pi)}{Y}\right)^{\frac{\alpha_h}{1-\phi_h}}}_{\tilde{\chi}_{hk}}$$

which is equation 5 in the main manuscript. By expressing 6 as a function of each country's GDP we have that

$$M^f_{tot} = \left[\sum_{j=1}^{N}\left(\frac{w_j}{w_k}\tau_{mjk}\right)^{1-\phi_h}\frac{\beta_{mj}}{\beta_{mk}}\frac{Y_j}{Y_k}\right]\left(\tilde{L}_k\right)\bar{m}^f_{hk} \tag{10}$$

where, $\left(\tilde{L}_k\right)\bar{m}^f_{hk} = M^f_{hk}$. Rearranging terms in the equation above we obtain that the variable χ_{hk} can be expressed as

$$\chi_{hk} = \left(\frac{M^f_{tot}}{M^f_k}\right)^{\frac{\alpha_h}{\phi_h-1}}$$

This last expression indicates that our variable capturing the gains from importing can be measured empirically by obtaining data on the volume of all intermediate inputs used at the firm level, M_{tot}^{f}, and the volume of domestic intermediate inputs used at the firm level M_{k}^{f}. We will obtain an estimation for this element in our empirical part.

2.3. Heterogeneity in quality and technology of intermediates

The goal of this subsection is to verify whether the predictions of our model change when considering a more complex and richer environment in which we allow for technological differences in the production of intermediates across countries and differences in quality across intermediate inputs.

Consider an alternative environment in which, to produce, a firm in the final good sector combines labour and intermediate inputs using a Cobb-Douglas technology as before. However, the expression for the intermediate input composite good is given by

$$m_{hn}^{f} = \left(\int_{\nu \epsilon \Lambda} \left(z_{hn} m_{hn}^{f}(\nu) \right)^{\frac{\phi_{h-1}}{\phi_h}} d\nu \right)^{\frac{\phi_h}{\phi_h - 1}}$$

where the parameter z_{hn} is a measure of the quality of the variety of intermediate input coming from country n. We assume that all intermediate input varieties coming from a specific location share the same quality. In addition, we assume that in each country intermediate inputs are produced using the following technology

$$m(\nu) = \zeta_n l_m$$

where ζ_n is the productivity of the intermediate input industry in country n. Note that there is symmetry across all varieties from the same source country. Analogous to the derivations in the previous section, we can obtain that a firm's production function can be rewritten as

$$q_{hk}^{f} = \varphi_h^{f} \left(l_{hk}^{f} \right)^{1-\alpha_h} \left(\tilde{M}_{tot}^{f} \right)^{\alpha_h} \left[\sum_{j=1}^{N} \left(\frac{w_j \zeta_k}{w_k \zeta j} \tau_{mjk} \right)^{1-\phi_h} \left(\frac{z_{hj}}{z_{hk}} \right)^{\phi_h - 1} \frac{\tilde{L}_j}{\tilde{L}_k} \right]^{\frac{\alpha_h}{\phi_h - 1}} \left(\tilde{L}_k \right)^{\frac{\alpha_h}{\phi_h - 1}} .$$

where $\tilde{M}_{tot}^{f} = \int_{\nu \epsilon \Lambda} \frac{\tilde{p}_{hmj}(\nu) \tilde{m}_{hj}^{f}(\nu)}{\tilde{p}_{hmk}} d\nu$, $\tilde{p}_{hmj}(\nu) = \frac{p_{hmj}(\nu)}{z_{hj}}$ is the price-adjusted quality of a variety coming from country j and $\tilde{m}_{hj}^{f}(\nu) = z_{hj} m_{hj}^{f}(\nu)$ is the quantity of intermediate inputs measured in quality units. The new definition for χ_{hk} is given by

$$\chi_{hk} = \left[\sum_{j=1}^{N} \left(\frac{w_j \zeta_k}{w_k \zeta_j} \tau_{mjk} \right)^{1-\phi_h} \left(\frac{z_{hj}}{z_{hk}} \right)^{\phi_h - 1} \frac{\tilde{L}_j}{\tilde{L}_k} \right]^{\frac{\alpha_h}{\phi_h - 1}} = \left(\frac{\tilde{M}_{tot}^{f}}{\tilde{M}_k^{f}} \right)^{\frac{\alpha_h}{\phi_h - 1}} = \left(\frac{M_{tot}^{f}}{M_k^{f}} \right)^{\frac{\alpha_h}{\phi_h - 1}} .$$

The definition of the variable χ_{hk} now depends on elements that control for technological differences across source intermediate input countries together with differences in the quality of intermediate inputs. This is the new term that captures the gain from importing. The rest of the results remain qualitatively equal. From this derivation, it becomes apparent that χ_{hk} can be retrieved from the data by using the statistic $\left(\frac{M_{tot}^{f}}{M_k^{f}} \right)^{\frac{\alpha_h}{\sigma_h - 1}}$. The last equality of the expression above comes because of the assumption of symmetry across all domestic varieties and the fact that both aggregates are deflated using the same price index. Consequently, the effect of changes in trade costs or market size on exports through imports can be perfectly captured by this element even if we assume a more realistic environment in which countries differ on the way they produce intermediates. Identically, the effects that changes in trade costs or market size have on exports via imports can be computed as described before.

2.4. Derivation of equations (13) and (14) of the main text.

In the previous section we have concluded that

$$\chi_{hk} = \left(\frac{M^f_{tot}}{M^f_k}\right)^{\frac{\alpha_h}{\phi_h - 1}} \quad .$$

Taking logs in the expression above and differentiating with respect to D_{kj} we have that

$$\frac{d \ln \chi_{hk}}{dD_{kj}} = \frac{\alpha_h}{\phi_h - 1} \frac{1}{\frac{M^f_{tot}}{M^f_k}} \frac{d\left(\frac{M^f_{tot}}{M^f_k}\right)}{dD_{kj}}$$

Note that

$$\frac{M^f_{tot}}{M^f_k} = \sum_{l=1}^{n} \frac{\tilde{L}_l}{\tilde{L}_k} \frac{p_{hml} m_{hml}}{p_{hmk} m_{hmk}}$$

and that

$$\frac{d \frac{M^f_{tot}}{M^f_k}}{dD_{kj}} = \frac{d \frac{\tilde{L}_j p_{hmj} m_{hmj}}{\tilde{L}_k p_{hmk} m_{hmk}}}{dD_{kj}} = \frac{d\left(\frac{M^f_j}{M^f_k}\right)}{dD_{kj}}$$

since $\frac{d \frac{\tilde{L}_l p_{hml} m_{hml}}{\tilde{L}_k p_{hmk} m_{hmk}}}{dD_{kj}} = 0$ for $l \neq j$.

Note also that $\frac{M^f_j}{M^f_k} = \frac{\tilde{L}_j p_{hmj} m_j}{\tilde{L}_k p_{hmk} m_{hmk}}$. Therefore,

$$\frac{d \ln \chi_{hk}}{dD_{kj}} = \frac{\alpha_h}{\phi_h - 1} \frac{1}{\frac{M^f_{tot}}{M^f_k}} \frac{d\left(\frac{M^f_j}{M^f_k}\right)}{dD_{kj}} \quad .$$

Multiplying and dividing by $\frac{M^f_j}{M^f_k}$ we obtain

$$\frac{d \ln \chi_{hk}}{dD_{kj}} = \frac{\alpha_h}{\phi_h - 1} \frac{\frac{M^f_j}{M^f_k}}{\frac{M^f_{tot}}{M^f_k}} \frac{d \ln\left(\frac{M^f_j}{M^f_k}\right)}{dD_{kj}} = \frac{\alpha_h}{\phi_h - 1} \frac{M^f_j}{M^f_{tot}} \frac{d \ln\left(\frac{M^f_j}{M^f_k}\right)}{dD_{kj}} \quad .$$

Multiplying in both sides by D_{kj} we obtain

$$\frac{d \ln \chi_{hk}}{d \ln D_{kj}} = \frac{\alpha_h}{\phi_h - 1} \frac{M^f_j}{M^f_{tot}} \frac{d \ln\left(\frac{M^f_j}{M^f_k}\right)}{d \ln D_{kj}}$$

which is equivalent to this expression

$$\rho^f_{D_j} = \frac{d \ln \chi_{hk}}{d \ln D_{kj}} = \frac{\alpha_h}{\phi_h - 1} * \frac{M^f_j}{\sum_{n=1}^{N} M^f_n} * \frac{d \ln\left(\frac{M^f_j}{M^f_k}\right)}{d \ln D_{kj}} \quad .$$

Table 1: Number of firms

Year	Active Firms	Exporters	Importers	Two-way traders
	(1)	(2)	(3)	(4)
2000	30,417	21,795	19,274	17,347
2001	30,029	21,896	19,466	17,647
2002	29,894	21,796	19,295	17,479
2003	28,921	21,147	18,379	16,691
2004	29,367	21,377	18,646	16,897
2005	30,190	21,751	18,975	17,211
2006	30,363	22,006	19,319	17,541

Note: The table reports, for each year, the number of manufacturing firms included in Micro.3.

An identical procedure allows us to conclude that

$$\rho^f_{Y_j} = \frac{d\ln \chi_{hk}}{d\ln Y_j} = \frac{\alpha_h}{\phi_h - 1} * \frac{M^f_j}{\sum_{n=1}^{N} M^f_n} * \frac{d\ln\left(\frac{M^f_j}{M^f_k}\right)}{d\ln Y_j} \quad .$$

3. Additional empirical analyses

3.1. Micro-level data

The empirical analysis combines two sources of data collected by the Italian Statistical Office (ISTAT): the Italian Foreign Trade Statistics (COE), and a firm-level accounting dataset (Micro 3).[4] ISTAT collects data on trade based on transactions. In compliance with the common framework defined by the European Union (EU), there are different requirements in order for a transaction to be recorded, depending on whether the importing country is an EU or NON-EU country, and on the value of the transaction.

As far as outside EU transactions are concerned, there is a good deal of homogeneity among member states as well as over time. In the Italian system the information is derived from the Single Administrative Document (SAD) which is compiled by operators for each individual transaction. Since the adoption of the Euro, Italy sets the threshold at 620 euro (or 1000 Kg), so that all transactions bigger than 620 euro (or 1000 Kg) are recorded. For all of these recorded extra-EU transactions, the COE data report complete information, that is, also information about the product quantity and value. Transactions within the EU are collected according to a different systems (Intrastat), where the thresholds on the annual value of transactions qualifying for a complete record are less homogeneous across EU member states, with direct consequences on the type of information reported in the data. In 2003 (the last year covered in the analysis), there are two cut-offs. If a firm has more than 200,000 euro of exports (based on previous year report), then a firm must fill the Intrastat document monthly. This implies that complete information about product is also available. Instead, if previous year export value falls in between 40,000 and 200,000 euro, the quarterly Intrastat file has to be filled, implying that only the amount of export is recorded, while information on the product is not. Firms with previous year exports below 40,000 euro are not required to report any information on trade flows. According to ISTAT, about one-third of the operators submitted monthly declarations, though covering about 98 percent of trade flows (http://www.coeweb.istat.it/default.htm). Thus, firms which do not appear in COE are either of this type (i.e. marginal exporters) or do not export at all.

The trade dataset is merged with Micro.3, which includes firm-level characteristics. After merging these two databases, we work with an unbalanced panel of about 48,179 manufacturing

[4]The database has been made available for work after careful screening to avoid disclosure of individual information

Table 2: Coverage of the dataset, 2003

	ALL FIRMS				EXPORTERS			IMPORTERS		
Sector	ASIA	Micro3	% Employees	% Sales	COE	Micro3	% Exports	COE	Micro3	% Imports
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
15	71,345	2,014	41.6	68.3	4,920	1,367	80.3	3,188	1,234	81.2
17	27,762	2,010	53.7	66.4	5,667	1,508	75.8	4,186	1,448	79.7
18	41,615	1,398	31.5	58.9	5,031	859	72.9	3,447	769	71.9
19	21,985	1,328	38.5	58.2	5,687	1,063	71.8	2,918	877	69.2
20	46,584	767	21.8	37.9	2,453	482	73.9	2,926	524	55.0
21	4,566	661	69.3	80.0	1,328	511	92.3	933	481	89.9
22	27,344	959	41.0	58.9	2,157	525	80.3	1,201	450	75.0
23	443	96	85.3	96.9	83	48	99.6	82	54	99.8
24	6,127	1,128	81.6	81.6	2,589	1,010	80.2	2,158	1,012	80.8
25	13,084	1,802	60.7	69.4	4,416	1,525	81.5	2,738	1,288	81.6
26	27,230	1,679	51.7	68.8	4,521	995	79.7	2,132	778	77.7
27	3,814	848	80.6	86.4	1,333	670	91.8	1,019	592	92.7
28	99,519	4,683	35.7	50.4	10,243	2,787	81.2	5,343	2,184	76.3
29	42,391	4,180	65.1	73.3	12,103	3,624	85.4	6,914	3,049	84.2
30	1,976	105	49.7	62.2	261	68	89.0	325	81	79.8
31	18,316	1,300	57.1	70.9	3,204	948	84.1	2,260	856	85.4
32	8,671	422	57.1	69.7	907	296	82.7	941	322	78.7
33	22,399	659	51.7	70.9	1,916	530	86.8	1,699	523	83.7
34	1,962	546	84.9	86.4	918	445	83.5	722	424	85.2
35	4,684	372	65.05	71.0	778	209	70.4	662	220	73.6
36	50,018	1,963	39.4	59.3	8,654	1,676	74.2	4,216	1,212	72.7
Total	541,836	28,921	50.9	70.1	79,170	21,147	81.5	50,011	18,379	82.5

Notes: The table reports, for 2003, the number of universe of active firms reported in ASIA (column 1), the number of active firms in Micro.3 (column 2), the coverage in terms of number of employees (column 3) and in terms of sales (column 4) of Micro.3. Columns 5-6 (8-9) report the number of exporters (importers) in COE and Micro.3, respectively. Column 7 (10) shows the coverage in terms of total exports (imports) of firms belonging to Micro.3. Sector definition according to 2-digit NACE manufacturing industries.

firms over the sample period. Column 1 of Table 1 presents the number of firms active in Micro.3.[5] Compared to the reference population of Italian manufacturing firms, our dataset covers only 6% of firms, but approximately 50% in terms of total employment and 70% for total sales. The main limitation of the sample is the mild over-representation of larger and more productive firms. As far as traders is concerned, columns 2 and 3 of Table 1 show the number of exporters and importers, respectively. The majority of traders are involved in both export and import activities, as reported in column 4 of the table. Two-way traders, which are more than 75% of trading firms, represent by far the largest share, while firms that only export or import are a relatively smaller fraction: around 10% are only-exporters and 15% are only-importers.[6]

To check the representativeness of our dataset with respect to the universe of Italian firms we use the ISTAT's archive of active firms, ASIA. The ASIA register covers the population of active Italian firms, irrespective of their trade status. It reports annual figures on firms' number of employees, total sales, sector of main activity and information about the geographical location of firms. Table 2 shows that the representativeness of our dataset, Micro.3, is quite satisfactory with respect to the universe. We report here 2003 data, but figures are comparable in the other years. As mentioned in the main text, although the dataset includes only about 6 percent of manufacturing in terms of number of firms, we cover about 50 percent of total number of employees and 70 percent of total sales. In terms of trade, our coverage is about 82 percent of total Italian exports and 83 percent of total Italian imports. We add here that these numbers are also basically stable across different industrial sectors.

The figures provided in Table 2 are explained by the well known abundance of micro and

[5] In our empirical analysis we are working with a slightly smaller sample due to missing values in some firms' variables reported in the balance sheet.

[6] The model focuses on an equilibrium in which firms are either non traders, importers who do not export or two-way traders. Indeed, the data confirms that the majority of firms are two-way traders while only a small fraction are firms that export without importing. In the analysis we add a robustness check where we exclude from the sample those firms that only export.

Table 3: Descriptive statistics: average firms' (log) total sales by sector, 2003

	ALL FIRMS		EXPORTERS		IMPORTERS	
	ln Sales					
	ASIA	Micro3	COE	Micro3	COE	Micro3
Sector	(1)	(2)	(3)	(4)	(5)	(6)
15	11.7	16.3	14.6	16.6	15.5	16.8
17	11.9	15.5	14.2	15.8	14.7	15.9
18	11.2	14.9	13.7	15.6	14.1	15.8
19	11.9	15.4	13.6	15.7	14.5	15.8
20	11.1	15.4	13.7	15.7	14.0	15.7
21	13.0	16.0	14.8	16.2	15.4	16.3
22	11.6	15.5	14.0	15.8	14.7	16.1
23	14.5	17.2	16.6	18.0	17.1	18.1
24	13.4	16.7	15.1	16.8	15.6	16.8
25	13.1	15.7	14.5	15.9	15.1	16.0
26	11.8	15.7	13.2	15.9	14.6	16.1
27	13.5	16.3	15.4	16.5	15.9	16.7
28	12.0	15.3	14.2	15.7	14.8	15.8
29	12.4	15.8	14.4	15.9	15.0	16.1
30	12.0	15.6	14.7	15.9	14.4	15.9
31	12.1	15.6	14.3	15.9	14.7	16.1
32	11.1	15.6	14.4	16.0	14.5	15.9
33	11.2	15.6	14.1	15.8	14.2	15.8
34	13.8	16.2	15.2	16.4	15.7	16.5
35	12.0	15.6	14.4	16.3	14.7	16.3
36	11.4	15.5	13.6	15.6	14.1	15.8

Notes: The table reports, for 2003, the average firms (log) sales by sector for: (1) all manufacturing firms; (2) firms in Micro.3; (3) all exporters; (4) exporters in Micro.3; (5) all importers; (6) importers in Micro.3. Sector definition according to 2-digit NACE manufacturing industries.

Table 4: Country-level variables

Variables	Proxies	Type of variable	Source
Y_{jt}	Gdp_{jt}	Continuous	World Bank
D_j	$Distance_j$	Continuous	CEPII
θ_{jt}	$Remoteness_{jt}$	Continuous	World Bank
F_j	$Market\ Costs_j$	Continuous	World Bank
Δ_j	$Trade\ Opening_{jt}$	Continuous	Fraser Institute

Note: The table reports the country-level variables used in the empirical analyses.

small firms in Italian manufacturing, together with the observation that medium-big firms are expected to account for the great bulk of overall export activities in the country, in line with a well established result in the literature. In agreement with this, Table 3 shows that (again for 2003 but valid across other sample years) the firms in our sample are on average slightly bigger, in terms of sales, than the population of manufacturing firms. At the same time, however, we do not observe big differences when we focus on exporting and importing firms: the average sales do not differ significantly between our sample and the population.

In addition to firm-level data, we complement the analysis with information on country characteristics. Table 4 lists the country-level characteristics used to proxy the variables in our empirical models. After selecting the destinations for which we have the information needed to carry out our analysis, we end up with a dataset including 134 countries. Table 5 reports the summary statistics for the country-level variables used in our empirical analysis.

3.2. Details for the production function estimation

In order to estimate firms' TFP by using a gross output production function in the presence of input endogeneity (i.e. firms choose inputs based on their observed productivity level), we rely on the method proposed by Gandhi et al. (2018). The proxy variable methods proposed by Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015) deal with this simultaneity (or transmission bias) problem mainly in the context of a value added

Table 5: Country variables: summary statistics

	Mean	SD	Min	25th Pct	75th Pct	Max
$\ln Distance_j$	8.28	0.90	6.20	7.59	9.05	9.83
$\ln GDP_{jt}$	23.94	2.07	19.14	22.39	25.51	29.99
$Trade\ Opening_{jt}$	7.334	1.43	3.49	6.27	8.54	9.98
$Market\ Costs_j$	-0.027	1.01	-1.57	-0.79	0.41	3.51
$\ln Remoteness_{jt}$	40.14	0.24	39.77	39.91	40.32	40.74

Note: The table reports the summary statistics for the country variables used in the empirical analysis. Statistics are computed on 134 countries.

production function (i.e. the intermediate input is not included in the estimated production function). However, Gandhi et al. (2017) has shown that a value-added production function can be constructed from an underlying gross output production function only under very restrictive hypotheses (such as the the linear in intermediate inputs Leontief specification) that are not compatible with our theoretical framework. Given that proxy variable methods are likely to suffer from identification issues when employed with a gross output production function,[7] we follow the suggestions of Ackerberg et al. (2015) and Gandhi et al. (2018) by exploiting the information contained in the first order condition of the firm's static profit maximization problem with respect to intermediate inputs.

In this section we show how to adapt the Gandhi et al. (2018) strategy to our theoretical setting that is characterized by the following production function (in which physical capital is added as an additional factor of production)

$$q_t^f = \left(m_t^f\right)^{\alpha} \left(k_t^f\right)^{\beta_k} \left(l_t^f\right)^{\beta_l} e^{\upsilon_t^f} \tag{11}$$

where, k_t^f, l_t^f represents respectively the amount of capital and labour used, and m_t^f is the intermediate composite input described, as in the main text, by the following expression

$$m_t^f = \left(\int\limits_{\nu \epsilon \Lambda} \left(m_t^f(\nu)\right)^{\frac{\phi-1}{\phi}} d\nu \right)^{\frac{\phi}{\phi-1}} .$$

Following Gandhi et al. (2018), the exponent of the element $e^{\upsilon_t^f}$ can be decomposed in two terms $\upsilon_t^f = \omega_t^f + \varepsilon_t^f$. The variable ω_t^f is the persistent part of the TFP that is known to the firm at time t, before the decisions on inputs have been taken. ε_t^f is an ex-post productivity shock not known by the firm at time t which is independent of the firms' information set at time t.[8] For simplicity, Gandhi et al. (2018) assumes that this productivity shock has zero mean. Following the proxy variable methods, in Gandhi et al. (2018) the law of motion of ω_t^f is represented by an exogenous first order Markov process $\omega_t^f = h(\omega_{t-1}^f) + \eta_t^f$ where η_t^f is also a productivity innovation that is independent of firms' choices at time $t-1$.

In the empirical analyses reported in the paper and in this appendix, the innate productivity parameter φ^f described in our theoretical model is proxied by the estimated $e^{\omega_t^f}$. We have made this choice because, with respect to ε_t^f, ω_t^f is closer to the static innate productivity concept used our theoretical model (which is a simplifying assumption common in the literature (Chaney, 2008)) due to its persistence and due to the fact that, similarly to φ^f, it is observed by firms at time t. However, all the empirical results are robust to using $e^{\omega_t^f+\varepsilon_t^f}$ as an alternative proxy for the innate firm productivity.

The method proposed by Gandhi et al. (2018) consists on inferring the firms' gross output

[7] On this point see Bond and Söderbom (2005), Ackerberg et al. (2015) and Gandhi et al. (2018).

[8] This element will also include possible measurement error to output and prices.

production function starting from the information captured by the first order conditions of the firms' profit maximization problem. In our theoretical model firms solve the following profit maximization problem

$$Max\ p_t^f \left(m_t^f\right)^{\alpha} \left(k_t^f\right)^{\beta_k} \left(l_t^f\right)^{\beta_l} e^{v_t^f} - w_t l_t^f - r_t k_t^f - \int\limits_{\nu \epsilon \Lambda} p_t^f(\nu) m_t^f(\nu)\, d\nu$$

$$s.t. \left(\int\limits_{\nu \epsilon \Lambda} \left(m_t^f(\nu)\right)^{\frac{\phi-1}{\phi}} d\nu \right)^{\frac{\phi}{\phi-1}} = m_t^f$$

where p_t^f denotes the price of the final variety. The first order condition associated with each of the varieties of the intermediate inputs is given by

$$p_t^f \left(\int\limits_{\nu \epsilon \Lambda} \left(m_t^f(\nu)\right)^{\frac{\phi-1}{\phi}} d\nu \right)^{\frac{\phi\alpha}{\phi-1}-1} \alpha \left(m_t^f(\nu)\right)^{\frac{-1}{\phi}} \left(k_t^f\right)^{\beta_k} \left(l_t^f\right)^{\beta_l} e^{\omega_t^f} E\left(e^{\varepsilon_t^f}\right) = p_t^f(\nu) \quad .$$

Multiplying both sides by $m_t^f(\nu)$ and integrating over ν we find that

$$p_t^f \left(\int\limits_{\nu \epsilon \Lambda} \left(m_t^f(\nu)\right)^{\frac{\phi-1}{\phi}} d\nu \right)^{\frac{\phi\alpha}{\phi-1}} \alpha \left(k_t^f\right)^{\beta_k} \left(l_t^f\right)^{\beta_l} e^{\omega_t^f} E\left(e^{\varepsilon_t^f}\right) = \int\limits_{\nu \epsilon \Lambda} p_t^f(\nu) m_t^f(\nu)\, d\nu \quad .$$

Note that we can rewrite this equation as

$$\alpha e^{-\varepsilon_t^f} E\left(e^{\varepsilon_t^f}\right) p_t^f q_t^f = \int\limits_{\nu \epsilon \Lambda} p_t^f(\nu) m_t^f(\nu)\, d\nu$$

and therefore, as in Gandhi et al. (2018), we obtain

$$S_t^f = \alpha e^{-\varepsilon_t^f} E\left(e^{\varepsilon_t^f}\right), \tag{12}$$

where S_t^f is the intermediate input share (i.e. $\frac{\int_{\nu\epsilon\Lambda} p_t^f(\nu) m_t^f(\nu) d\nu}{p_t^f q_t^f}$), which is observable by the econometrician.

The first stage of the Gandhi et al. (2018) concentrates on obtaining an estimate of $\alpha = \frac{\partial \ln q_t^f}{\partial \ln m_t^f}$. Taking logs in 12 and defining $s_t^f = \ln S_t^f$, we obtain the following share equation (from which the productivity term ω_t^f inducing transmission bias is absent)

$$s_t^f = \underbrace{\ln \alpha + \ln E\left(e^{\varepsilon_t^f}\right)}_{\ln D} - \varepsilon_t^f \quad .$$

To identify $\frac{\partial \ln q_t^f}{\partial \ln m_t^f}$, we follow Gandhi et al. (2018) by obtaining an estimate of the regression function $\widehat{\ln D}$ defined by the above share equation using Non Linear Least Squares.[9] In this

[9]Gandhi et al. (2018) does not assume any functional form for the production function and approximate D with a second order polynomial in k, l and m. In our model a Cobb-Douglas production function is assumed, and given the separability of the production function, we can conclude that the intermediate input share is a constant, given by the exponent of the intermediate input composite good in the production function, α. In additional robustness checks not reported in this appendix, we have applied their non-parametric specification

way we are also able to retrieve an estimate of the residuals $\hat{\varepsilon}_t^f$ corresponding to the ex-post shocks and of $E\left(e^{\varepsilon_t^f}\right)$, $\frac{1}{N\times T}\sum_{f,t} e^{\hat{\varepsilon}_t^f}$. Therefore, it is possible to obtain

$$\widehat{\frac{\partial \ln q_t^f}{\partial \ln m_t^f}} = \hat{\alpha} = \frac{\exp(\widehat{\ln D})}{\frac{1}{N\times T}\sum_{f,t}\exp(\widehat{\ln D} - s_t^f)} \quad .$$

Once obtained the above estimate of α, Gandhi et al. (2018) notes that, by integrating the partial differential equation defined by the intermediate input elasticity, it is possible to obtain the portion of the production function related to the intermediate input

$$\int \frac{\partial \ln q_t^f}{\partial \ln m_t^f} d\ln m_t^f = \ln q_t^f + h(k_t^f, l_t^f) \tag{13}$$

where $h(k_t^f, l_t^f)$ is the constant of integration. Taking logs in equation 11, substracting equation 13 and rearranging terms we find that

$$\ln q_t^f - \varepsilon_t^f - \alpha \ln m_t^f = \omega_t^f - h(k_t^f, l_t^f) \quad . \tag{14}$$

In section 2.2 of this document we show that

$$\left(m_t^f\right)^\alpha = \left(M_{tot,t}^f\right)^\alpha \left[\sum_{j=1}^N \left(\frac{w_j}{w}\tau_{mj}\right)^{1-\phi} \frac{\tilde{L}_l}{\tilde{L}}\right]^{\frac{\alpha}{\phi-1}} \left(\tilde{L}\right)^{\frac{\alpha}{\phi-1}} \tag{15}$$

where $M_{tot,t}^f$ represents the total volume of intermediate inputs used by the firm including both domestic and foreign inputs and

$$\left[\sum_{j=1}^N \left(\frac{w_j}{w}\tau_{mj}\right)^{1-\phi} \frac{\tilde{L}_j}{\tilde{L}}\right]^{\frac{\alpha}{\phi-1}} = \left(\frac{M_{tot,t}^f}{M_t^f}\right)^{\frac{\alpha}{\phi-1}} .$$

where M_t^f represents the total volume of domestic intermediate inputs. Inserting the expression 15 for $(m_t^f)^\alpha$ in the production function (and taking the natural logarithm of it) we find

$$\alpha \ln m_t^f = \alpha \ln M_{tot,t}^f + \frac{\alpha}{\phi - 1}\ln\left(\frac{M_{tot,t}^f}{M_t^f}\right) + \frac{\alpha}{\phi-1}\ln\tilde{L} \quad .$$

Plugging this expression in equation in 14 we obtain

$$\ln q_t^f - \varepsilon_t^f - \alpha\left(\ln M_{tot,t}^f + \frac{1}{\phi-1}\ln\left(\frac{M_{tot,t}^f}{M_t^f}\right) + \frac{1}{\phi-1}\ln\tilde{L}\right) = \omega_t^f - h(k_t^f, l_t^f) \quad .$$

Rearranging terms,

$$\ln q_t^f - \varepsilon_t^f - \alpha \ln M_{tot,t}^f = \omega_t^f + \frac{\alpha}{\phi-1}\ln\left(\frac{M_{tot,t}^f}{M_t^f}\right) + \frac{\alpha}{\phi-1}\ln\tilde{L} - h(k_t^f, l_t^f) \quad .$$

Note that $\frac{\alpha}{\phi-1}\ln\left(\frac{M_{tot,t}^f}{M_t^f}\right)$ represents, as described in the appendix, the contribution of

and we have verified that the sectoral estimated average elasticities are very similar to those obtained for the Cobb-Douglas case.

imported intermediate inputs to the firm's TFP.[10] Given that our model suggests a Cobb-Douglas production function, we will consider that our constant of integration will be given by $-h(k_t^f, l_t^f) = \beta_k \ln k_t^f + \beta_l \ln l_t^f$. Therefore, following Gandhi et al. (2018) we need to estimate the following equation

$$\mathcal{Y}_t^f \equiv \ln q_t^f - \varepsilon_t^f - \alpha \ln M_{tot,t}^f = \omega_t^f + \frac{\alpha}{\phi - 1} \ln \left(\frac{M_{tot,t}^f}{M_t^f} \right) + \frac{\alpha}{\phi - 1} \ln \tilde{L} + \beta_k \ln k_t^f + \beta_l \ln l_t^f$$

where $\mathcal{Y}_t^f$ is a function of an observable, $M_{tot,t}^f$, and of other two elements which can be obtained from the share regression estimation: the intermediate input elasticity and the ex-post shock.

At this point it is important to note that ω_t^f can be rewritten as a function of $\mathcal{Y}_t^f$ and of other observable variables multiplied by the corresponding elasticities

$$\omega_t^f = \mathcal{Y}_t^f - \frac{\alpha}{\phi - 1} \ln \left(\frac{M_{tot,t}^f}{M_t^f} \right) - \frac{\alpha}{\phi - 1} \ln \tilde{L} - \beta_k \ln k_t^f - \beta_l \ln l_t^f \quad .$$

By following the the proxy variable literature, Gandhi et al. (2018) exploit the dynamic Markovian structure on productivity to express the innovation of productivity, η_t^f, as a function of the remaining parameters $\eta_t^f \left(\beta_l, \beta_k, \frac{\alpha}{\phi-1} \right)$ to be estimated by regressing $\omega_t^f \left(\beta_l, \beta_k, \frac{\alpha}{\phi-1} \right)$ on $g \left(\omega_{t-1}^f \left(\beta_l, \beta_k, \frac{\alpha}{\phi-1} \right) \right)$. This final step proceeds with an iterative Generalised Method of Moments (GMM). The parameters of interest $\left(\beta_l, \beta_k, \frac{\alpha}{\phi-1} \right)$ are identified relying on the following moment conditions

$$E \left\{ \eta_t^f \left(\beta_l, \beta_k, \frac{\alpha}{\phi - 1} \right) \begin{pmatrix} \ln \left(l_t^f \right) \\ \ln \left(k_t^f \right) \\ \ln \left(\frac{M_{tot,t-1}^f}{M_{t-1}^f} \right) \end{pmatrix} \right\} = 0 \tag{16}$$

where both capital (as standard) and labor (due to the strong Italian employment protection legislation) are considered as state variables. We expect current intermediate input sourcing choices to be correlated with shocks to productivity and, therefore, we rely on lagged choices (following the reasoning of Kasahara and Rodrigue (2008) we consider past import status as an additional state variable).

In our application, in order to take into account the possibility of the existence of an endogenous productivity process characterized by a dynamic learning by importing effect, we follow Kasahara and Rodrigue (2008) and, for the main results of the paper, we let the dynamics of productivity to potentially depend on imported inputs

$$\omega_t^f = \underbrace{g \left(\omega_{t-1}^f, \ln \frac{M_{tot,t-1}^f}{M_{t-1}^f} \right)}_{\text{expected productivity}} + \underbrace{\left(\eta_t^f \right)}_{\text{productivity shock}} \quad . \tag{17}$$

Our baseline empirical implementation of the law of motion for productivity allows the persistence in productivity and the effect of importing on future productivity (i.e., the learning

[10] The element $\frac{\alpha}{\phi-1} \ln \tilde{L}$ reflects the gains from variety in the use of domestic intermediates. This element is common to all firms belonging to a given sector. As it is well known, a constant in the production function and mean productivity cannot be separately identified.

Table 6: Learning by importing and Learning by exporting

Dep. Var.	$\widehat{\omega}_t^f$			
Indep. Var.	$\widehat{\omega}_{t-1}^f$	$\ln \frac{M_{tot,t-1}^f}{M_{t-1}^f}$	$\ln\left(\frac{\text{Total Sales}}{\text{Domestic Sales}}\right)_{t-1}^f$	N.Obs
Food, Beverages and Tobacco	0.956***	0.015*	0.001	9875
Textiles and Apparel	0.964***	0.086***	0.006***	16570
Hide and Leather	0.935***	0.075***	0.005***	6517
Wood and Cork	0.919***	0.045***	0.001	3751
Pulp and Paper	0.938***	0.029***	0.000	3350
Printing and Publishing	0.954***	0.040***	0.000	4847
Coke and Chemical products	0.952***	0.019***	0.001	6334
Rubber and Plastics	0.943***	0.038***	0.001	9258
Processing of non-metallic minerals	0.928***	0.030***	0.000	8381
Basic Metals	0.948***	0.019***	0.001	4268
Fabricated Metal Products	0.937***	0.063***	0.006***	23745
Machinery and Equipment	0.920***	0.042***	0.004***	21647
Electrical and Optical Equipment	0.931***	0.041***	0.005**	12297
Motor Vehicles and Trailers	0.920***	0.028***	0.009***	2855
Other Transport Equipment	0.880***	0.067***	0.013**	1726
Other manufacturing industries	0.938***	0.041***	0.001	9876

Note: The table reports the estimated coefficients for the law of motion of ω^f by sector using data on 2000-2006. Asterisks denote significance levels obtained with bootstrapped standard errors (500 replications) (***:p<1%; **: p<5%; *: p<10%).

by importing effect) to depend on a firm's productivity level

$$g_{bas}(\cdot) = \sum_{m=1}^{3} \psi_m (\omega_{t-1}^f)^m + \psi_4 \ln\left(\frac{M_{tot,t-1}^f}{M_{t-1}^f}\right) + \sum_{m=1}^{3} \psi_{4+m} \left(\omega_{t-1}^f\right)^m \times \ln\left(\frac{M_{tot,t-1}^f}{M_{t-1}^f}\right) \tag{18}$$

In the robustness checks of the paper, the above baseline empirical implementation of the law of motion for productivity is augmented for allowing the dynamics of productivity to potentially depend on exporting behaviour (as argued by De Loecker (2013)) which is proxied by $\ln E_{t-1}^f = \ln(\text{Total Sales}/\text{Domestic Sales})_{t-1}^f$

$$\begin{aligned} g_{rob}(\cdot) = &\sum_{m=1}^{3} \psi_m (\omega_{t-1}^f)^m + \psi_4 \ln\left(\frac{M_{tot,t-1}^f}{M_{t-1}^f}\right) + \sum_{m=1}^{3} \psi_{4+m} \left(\omega_{t-1}^f\right)^m \times \ln\left(\frac{M_{tot,t-1}^f}{M_{t-1}^f}\right) + \\ &\psi_8 \ln E_{t-1}^f + \sum_{m=1}^{3} \psi_{8+m} \left(\omega_{t-1}^f\right)^m \times \ln E_{t-1}^f \end{aligned} \tag{19}$$

For the sake of brevity and facility of interpretation, in Table 6 of this Appendix we report the estimated persistence ($\hat{\varrho}_0$), learning by importing ($\hat{\varrho}_1$) and learning by exporting ($\hat{\varrho}_2$) coefficients for the following specification[11]

$$g_{app}(\cdot) = \varrho_0 \omega_{t-1}^f + \varrho_1 \ln\left(\frac{M_{tot,t-1}^f}{M_{t-1}^f}\right) + \varrho_1 \ln E_{t-1}^f. \tag{20}$$

The results suggest that the learning by importing effects are positive and significant, while the learning by exporting effects are of much lower magnitude and statistically significant only for about half of the sectors.

[11]The estimated parameters $\left(\hat{\beta}_l, \hat{\beta}_k, \frac{\hat{\alpha}}{\phi-1}\right)$ obtained with the alternative law of motion of productivity $g_{rob}(\cdot)$ and $g_{rob}(\cdot)$ are very similar to those reported in the main text, which are obtained using $g_{bas}(\cdot)$.

Table 7: Firms' exports extensive and intensive margin by country. OLS regression

Dep. Var.	$ExportStatus^f_{jt}$		$\ln Export^f_{jt}$		$Export^f_{jt}$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\ln \widehat{\varphi}^f_t$	0.017***	0.017***	0.695***	0.695***	1.088***	1.064***
	(0.001)	(0.001)	(0.026)	(0.027)	(0.088)	(0.089)
$\ln \widehat{\chi}^f_t$	0.034***	0.034***	1.157***	1.155***	2.206***	2.174***
	(0.004)	(0.004)	(0.083)	(0.084)	(0.177)	(0.177)
$\ln Gdp_{jt}$	0.039***	0.037***	0.497***	0.463***	0.853***	0.832***
	(0.000)	(0.000)	(0.003)	(0.003)	(0.014)	(0.022)
$\ln Distance_j$	-0.063***	-0.056***	-0.599***	-0.470***	-1.105***	-0.964***
	(0.000)	(0.000)	(0.006)	(0.007)	(0.036)	(0.059)
$Trade\ Opening_{jt}$	0.009***	0.007***	0.045***	0.039***	0.134***	0.050***
	(0.000)	(0.000)	(0.002)	(0.002)	(0.013)	(0.018)
$\ln Remoteness_{jt}$	0.025***	0.059***	0.684***	0.391***	0.998***	1.009***
	(0.001)	(0.001)	(0.021)	(0.027)	(0.140)	(0.190)
$Market\ Costs_j$	-0.006***	-0.008***			0.111***	-0.248***
	(0.000)	(0.000)			(0.036)	(0.048)
Year FE	Yes	-	Yes	-	Yes	-
Year*Area FE	-	Yes	-	Yes	-	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
N	18,192,983	18,192,983	1,959,454	1,959,454	15,025,785	15,025,785
adj. R^2	0.291	0.304	0.306	0.322	-	-

Note: The table reports regressions using data on 2001-2006 for the extensive margin (columns 1-2), the intensive margin (columns 3-4), the Poisson model (columns 5-6). The dependent variable used is reported at the top of the columns. All the regressions include a constant term. Robust standard errors clustered at the firm and country-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

3.3. The extensive and intensive margins of exports

Equation 11 and 12 of the paper form the underpinning of our estimations. Precisely, equation 11 predicts that the country-by-country export decision depends on a firm's innate productivity (φ), foreign market size (Y), the multilateral resistance term (θ), variable trade costs (D and Δ), fixed trade costs (F), the contribution to TFP of importing intermediate inputs χ and other variables which are constant across firms. Similarly, all these elements, except the fixed trade costs, enter in the individual export value decision. The fixed costs, once paid, do not influence an exporter's foreign sales.

Columns 1 and 2 of Table 7 reports the estimation results of equation 11 in the paper, by using a simple OLS approach. We start in column 1 of Table 7 by reporting the results of a model which controls for firm and year fixed effects. Since it might be that there are time-varying effects common to countries belonging to the same geographical area, in column 2 of Table 7 we replicate the analysis by including year-area fixed effects. All the results are confirmed in both specifications and they are in line with the instrumental variable approach reported in the paper. It is instructive to notice that the IV coefficients reported in the paper are larger than the OLS estimates. A similar increase in the coefficient is observed in Mion and Zhu (2013); Bas and Strauss-Kahn (2014); Feng et al. (2016). Columns 3 and 4 of Table 7 report the results of equation 12 of the paper, by using simple OLS. As for the export decision equation, we run the regression controlling for year dummies (column 3) and then taking into account year-area fixed effects (column 4). Column 3 displays the results of our baseline specification. The results in column 4 include the control for the year-area dummies and they are qualitatively similar to those reported in column 3. Finally, columns 5 and 6 of Table 7 consider the estimation of the intensive margin equation in its multiplicative form with a pseudo-maximum-likelihood technique. Looking at the results we can conclude that the main message with respect to the previous specifications does not change.

To identify the causal effect of the TFP related to imported inputs on firms' export activities we apply an instrumental variable approach, using two different instruments described in the paper. Table 8 reports the estimate of the first stage of the IV estimation results for both

Table 8: First Stage of the IV estimation

Dep. Var.	$\ln \widehat{\varphi}_t^f$					
	Extensive Margin			Intensive Margin		
	(1)	(2)	(3)	(4)	(5)	(6)
	IV_{GDP_f}	$IV^P_{EUM_f}$	IV_{GDP_f},$IV^P_{EUM_f}$	IV_{GDP_f}	$IV^P_{EUM_f}$	IV_{GDP_f},$IV^P_{EUM_f}$
IV_{GDP_f}	0.612***		0.506***	0.469***		0.388***
	(0.038)		(0.042)	(0.035)		(0.037)
$IV^P_{EUM_f}$		0.639***	0.475***		0.459***	0.351***
		(0.042)	(0.046)		(0.040)	(0.044)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year*Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Other second-stage exogenous variables	Yes	Yes	Yes	Yes	Yes	Yes
N	8,582,803	8,504,615	8,225,922	1,435,557	1,426,550	1,389,028

Note: The table reports the results of the first state of IV regressions. Columns 1-2-3 report the first stage for columns 1-4-7 of Table 2 in the paper. Columns 4-5-6 report the first stage for columns 2-5-8 of Table 2 in the paper. Robust standard errors clustered at the firm and destination-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

Table 9: Estimating the coefficients for the country-level variables

Dep. Var.	Estimated fixed effects extensive margin		Estimated fixed effects intensive margin	
	(1)	(2)	(3)	(4)
$\ln GDP_{jt}$	0.060***	0.060***	0.387***	0.518***
	(0.005)	(0.005)	(0.041)	(0.040)
$\ln Distance_j$	-0.098***	-0.099***	-0.345***	-0.613***
	(0.012)	(0.013)	(0.079)	(0.087)
$Trade\ Opening_{jt}$	0.013***	0.013***	0.023	0.048*
	(0.004)	(0.004)	(0.022)	(0.029)
$\ln Remoteness_{jt}$	0.054	0.058	0.200	0.728***
	(0.035)	(0.041)	(0.214)	(0.256)
$Market\ Costs_j$	-0.013*	-0.013*		
	(0.007)	(0.007)		
Year	-	Yes	-	Yes
Firm FE	-	Yes	-	Yes
Cluster	Country	Country and Firm	Country	Country and Firm
N	7,940,616	7,940,616	1,256,783	1,256,783
adj. R^2	0.241	0.453	0.088	0.497

Note: The table reports regressions using data on 2000-2006. The dependent variables are the estimated fixed effects obtained in column 4 of Table 3 for the extensive margin (columns 1-2) and in column 4 of Table 4 for the intensive margin (columns 3-4). Clustered robust standard errors are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

the extensive margin (columns 1-3) and the intensive margin (column 4-6). As expected, both instruments are positively correlated with $\ln \widehat{\varphi}_t^f$. Indeed, the higher is the (weighted) gross domestic product, proxing for the number of available varieties of foreign intermediate inputs, the greater is the TFP-enhancing effect of imported intermediate inputs. Similarly, the variation of aggregate imports at the product level for similar developed countries is a good predictor of changes in the TFP enhancing effect of firm-level imports.

In column 4 of Tables 3 and 4 of the paper, we estimate the extensive and intensive margins equations by including firm-destination and year-destination fixed-effects. In this case, identification of our key variable relies only on variations over time of a firm's exports to the same destination, controlling for time variant and time invariant country characteristics. This specification is appealing because our proxies for F_j and θ_{jt} can be somewhat controversial. In Table 9 we propose an alternative model where we estimate the coefficient of Y_{jt}, θ_{jt}, D_j, Δ_{jt}, F_j in equations (11)-(12) by regressing the sum of the estimated year-destination fixed effects and firm-destination fixed effects on Y_{jt}, θ_{jt}, D_j, Δ_{jt}, F_j. Columns 1 and 3 of Table 9 report the results of a simple OLS regression, while columns 2 and 4 add year and firm-fixed effects. For both the extensive and the intensive margins, the magnitude of the estimated coefficients

Table 10: Estimating the coefficients for the country-level variables: alternative specification

Dep. Var.	$ExportStatus^f_{jt}$	Estimated year-destination FE	$lnExport^f_{jt}$	Estimated year-destination FE
	(1)	(2)	(3)	(4)
$\ln \widehat{\varphi}^f_t$	0.105***		1.0191***	
	(0.012)		(0.121)	
$\ln \widehat{\chi}^f_t$	0.572***		5.361***	
	(0.056)		(0.846)	
$\ln GDP_{jt}$		0.060***		0.538***
		(0.005)		(0.040)
$\ln Distance_j$		-0.096***		-0.604***
		(0.012)		(0.076)
$Trade\ Opening_{jt}$		0.012***		0.075*
		(0.004)		(0.044)
$\ln Remoteness_{jt}$		0.049		0.381
		(0.035)		(0.274)
$Market\ Costs_j$		-0.013*		
		(0.007)		
Firm	Yes	-	Yes	-
Year-Destination FE	Yes	-	Yes	-
Cluster	Country and Firm	Country	Country and Firm	Country
N	8,225,922	751	1,389,028	153
adj. R^2	0.372	0.832	0.342	0.879

Note: The table reports regressions using data on 2000-2006 for the extensive margin (columns 1-2) and the intensive margins (3-4). In columns 1 and 3 the instrumental variable are IV_{GDP_f} and $IV^P_{EUM_f}$. Clustered robust standard errors are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

of our country-level variables are comparable with those obtained in our baseline specification, in columns 7 and 8 of Table 2 of the paper.

Because the estimated coefficients used as dependent variable in Table 9 include both year-destination and firm-destination fixed effects, we also provide evidence of a different specification where we run a first-step including only firm and year-destination fixed effects. We then regress the estimated year-destination fixed effects on the country-level variable. Table 10 shows the results for both the extensive margin (columns 1-2) and the intensive margins (3-4). Results do not change with respect to those observed in Table 9.

In the paper, to take into account firms' unobserved heterogeneity, we estimate the intensive margin equation using the level of exports as dependent variable by employing a conditional (firm) fixed-effects Poisson model, which is appropriate for nonlinear models such as the gravity equation (Silva and Tenreyro, 2006). The main advantage of the Poisson estimator is that it naturally includes observations for which the observed trade value is zero, that is it takes into account the extensive and the intensive margins at the same time. Such observations are dropped from the OLS model because the logarithm of zero is undefined.

As an additional robustness check, we estimate a two-stage procedure in the spirit of Heckman's method to control for possible selection bias. We include the polynomials of the predicted value of $ExportStatus^f_{jt}$, obtained after estimating equation 11 and replacing any predicted value outside the 0-1 range so to bound the value between 0 and 1, into equation 12. Table 11 shows the results using as instrument the IV_{GDP_f} (column 1), the $IV^P_{EUM_f}$ (column 2), and both of them (column 3). The estimated coefficients remain robust to this alternative model.

Table 11: Firms' exports intensive margin: controlling for selection effect

Dep. Var.	$\ln Export^f_{jt}$		
	(1)	(2)	(3)
	IV_{GDP_f}	$IV^P_{EUM_f}$	IV_{GDP_f},$IV^P_{EUM_f}$
$\ln \widehat{\varphi}^f_t$	1.383***	1.495***	1.395***
	(0.170)	(0.216)	(0.169)
$\ln \widehat{\chi}^f_t$	6.869***	7.909***	7.023***
	(1.089)	(1.518)	(1.070)
$\ln GDP_{jt}$	0.329***	0.328***	0.330***
	(0.025)	(0.025)	(0.025)
$\ln Distance_j$	-0.277***	-0.275***	-0.279***
	(0.051)	(0.051)	(0.051)
$Trade\ Opening_{jt}$	-0.000	-0.001	0.001
	(0.019)	(0.019)	(0.019)
$\ln Remoteness_{jt}$	0.180	0.176	0.177
	(0.180)	(0.180)	(0.180)
$Predicted\ Value^f_{jt}$	3.316***	3.317***	3.329***
	(0.082)	(0.081)	(0.081)
Year*Area FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
N	1,395,637	1,400,100	1,361,554
adj. R^2	0.196	0.194	0.197
Underidentification stat.	38.941	37.744	41.796
(p-value)	0.000	0.000	0.000
Weak identification stat.	180.127	135.463	116.497
Hansen J stat.			0.662
(p-value)			0.416

Note: The table reports regressions using data on 2000-2006. Robust standard errors clustered at the firm and destination-level are reported in parenthesis below the coefficients. Asterisks denote significance levels (***:p<1%; **: p<5%; *: p<10%).

References

Ackerberg, D. A., K. Caves, and G. Frazer (2015): "Identification Properties of Recent Production Function Estimator," *Econometrica*, 83, 2411–2451.

Bas, M. and V. Strauss-Kahn (2014): "Does importing more inputs raise exports? Firm level evidence from France," *Review of World Economics*, 150, 241–475.

Bond, S. and M. Söderbom (2005): "Adjustment costs and the identification of Cobb Douglas production functions," Tech. rep., IFS Working Papers, Institute for Fiscal Studies (IFS).

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

De Loecker, J. (2013): "Detecting Learning by Exporting," *American Economic Journal: Microeconomics*, 5, 1–21.

Feng, L., Z. Li, and D. L. Swenson (2016): "The Connection between Imported Intermediate Inputs and Exports: Evidence from Chinese Firms," *Journal of International Economics*, 101, 86–101.

Gandhi, A., S. Navarro, and D. Rivers (2018): "On the Identification of Gross Output Production Functions," University of Western Ontario, Centre for Human Capital and Productivity (CHCP) Working Papers 20181, University of Western Ontario, Centre for Human Capital and Productivity (CHCP).

Gandhi, A., S. Navarro, D. Rivers, et al. (2017): "How heterogeneous is productivity? A comparison of gross output and value added," *Journal of Political Economy, forthcoming.*

Kasahara, H. and J. Rodrigue (2008): "Does the use of imported intermediates increase productivity? Plant-level evidence," *Journal of Development Economics*, 87, 106–118.

Levinsohn, J. and A. Petrin (2003): "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70, 317–41.

Mion, G. and L. Zhu (2013): "Import competition from and offshoring to China: A curse or blessing for firms?" *Journal of International Economics*, 89, 202–215.

Olley, G. S. and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Silva, J. M. C. S. and S. Tenreyro (2006): "The Log of Gravity," *The Review of Economics and Statistics*, 88, 641–658.