Abstract

Using an exhaustive dataset on firm-to-firm sales in Belgium, we examine how firms choose their local sourcing strategy based on geography and their productivity. We document that most firms trade only locally, and that they follow a predictable pecking order among source sectors. These empirical findings motivate and guide the development of a model with endogenous choice of tasks produced in-house. Consistent with the model, we show that the probability of a trade relationship increases with the efficiency of both the supplier and the buyer and decreases with the distance between them. Finally, we run a counterfactual exercise to assess to what extent local trade frictions shape the productivity distribution.
1 Introduction

Dramatic advances in communication, information, and transportation technologies have led firms to reassess the traditional way to organize themselves and have contributed to a fragmentation of production processes. Firms are increasingly focusing on their core competencies and hiring outside companies to provide tasks that were previously done in-house (see Weil, 2014, for a qualitative survey). Labor cost savings are an important reason why firms choose to contract out these tasks (Goldschmidt and Schmieder, 2016). This behaviour is widely spread in the economy and is typically related to the sourcing of non-specific material inputs or services. For instance, according to the Belgian Structure of Earning Survey 2010, 77% of Belgian firms were sourcing (part of) their accounting services, 72% some office administrative services, 57% their legal services and 38% their cleaning services. In 2015, these shares were respectively of 80%, 76%, 66% and 51%. These inputs do not represent a large fraction of the total costs of the firm, as on average each category accounts for 1 or 2% of total input consumption, but most firms tend to prefer to buy them from other firms instead of producing them in house.

In this paper, we stress the role of the extensive margin of local sourcing on production efficiency. Figure 1 illustrates that the TFP premium at the firm level is increasing in the number of domestic suppliers. A firm with 50 suppliers is about 0.5 log point more productive than a firm operating in the same sector but which only source inputs from one domestic supplier. This result extends the evidence on global sourcing that the relative size advantage of manufacturing firms is increasing in the number of countries from which they source (Antrás et al., 2017). These efficiency advantages are suggestive of fixed cost of sourcing at the firm-to-firm level, which limits the ability of less efficient firms to select into sourcing from a large set of specialized suppliers.

The goal of this paper is to document local sourcing by using an exhaustive dataset on firms’ buyer-seller linkages in Belgium. In contrast to the Structure of Earning Survey which documents the sourcing strategy of a few thousand firms, this dataset allows us to characterize the local sourcing strategy followed by the universe of Belgian firms. First, we provide descriptive statistics on the distance between the supplier and the buyer by stressing the fact that most firms tend to trade with local partners. Perhaps more surprisingly, selection into local sourcing exhibits complementarities across sectors of activities. As a result, firms follow a predictable pecking order in their sourcing decision. Even though most firms buys inputs from suppliers operating in the wholesale, retail, transportation or network industries (energy provider or telecommunications)\footnote{In the 2010 Structure of Earning survey, around 90% of the sampled firms were reporting the consumption of electricity and telecommunication services.}, inputs from services sectors and, to a lesser extent, from the manufacturing sectors are crucial in explaining the heterogeneity in the extensive margins of local sourcing. Legal and accounting activities, office administrative support activities and computer programming are sectors ranked 1 to 3.
To guide the empirical analysis, we modify the global sourcing model of Antrás et al. (2017) to allow for the possibility that firms choose endogenously the amount of tasks performed in-house in order to produce a final good or service. We exploit the complementarity mechanism herein to rationalize the sourcing decisions made by firms. This simple model puts us in a position to derive some testable assumptions on the sourcing decisions. More efficient and less distant suppliers are more likely to be chosen. The share of tasks is increasing in the supplier’s productivity and decreasing both in the geographic distance between the supplier and the buyer and in the buyer’s productivity.

The results obtained in the empirical part are basically in line with the predictions of the model. These results are robust to alternative specifications involving different level of localisation dummies or firm fixed effects. From a sectoral analysis, we show that the impact of distance, even if it is always sizeable, varies strongly across sectors both within the manufacturing or the services sectors but do not differ significantly across those two broad sectors.

Our paper relates to three strands of literature. This paper is naturally related to a growing literature on the determinants of domestic (Bernard et al., 2017, Furusawa et al., 2017) and foreign sourcing (Amiti and Koning, 2007, Goldberg et al., 2010, Halpern et al., 2015, Bøler et

Figure 1 – TFP Premia and domestic sourcing

Note: To construct the blue line, we regress the log TFP on cumulative dummies for the minimum number of domestic suppliers from which the firm sources, and sector controls. To construct the red line, we regress the TFP on cumulative dummies for the minimum number of domestic suppliers from which the firm sources, cumulative dummies for the minimum number of foreign countries from which the firm sources, and sector controls. These premia are robust to controlling for the zip code, and thus do not merely capture the fact that more productive firms are benefiting from agglomeration effects.
al., 2015, Antrás et al., 2017) and the impact on firms.

Second, our paper contributes to the growing literature on domestic production networks (Lim, 2015). Oberfield (2013) develops a general-equilibrium theory of the structure of production based buyer’s optimal choices of suppliers to study how the endogenous network formation shapes the productivity and organization of production.

Third, our paper relates to literature on economic geography. A large body of empirical literature finds that productivity increases with agglomeration effects (for a survey, see Rosenthal and Strange, 2004), but there are only rare attempts to distinguish the various channels behind agglomeration economies. These have been traditionally grouped into the Marshallian three broad families of agglomeration mechanisms: labor pooling, knowledge spillovers, and input-output linkages. These three potential channels have not been easy to disentangle from an empirical perspective. With a full knowledge of the sourcing decision made by firms, our paper intends to focus specifically on the third channel. To do so, we construct an indicator of firm connectivity and relate it to the firm location. Firms located in places where connectivity to suppliers is high exhibit higher productivity levels. Our estimates suggest that total factor productivity in Belgium would be on average 5.5% higher if all firms were as well connected as in the best connected area in Belgium.

Belgium provides a particularly interesting setting to conduct such a counterfactual exercise, since we expect trade frictions to be at a lower bound. Belgium is a bilingual country, so admittedly cultural barriers to trade between firms located in the Flemish or the Walloon region may be at play. In the empirical section, we will therefore introduce a control in our regression to take this barrier into account. Otherwise, with more than 780,000 firms within a geographic area of 30,000 km², Belgium has one of the highest firm density in Europe. We consider that, because of its small size, Belgium deserves specific attention. Distance between firms is at most 277 km. Belgium has a very dense transportation infrastructure (155,000 km of roads, 3,500 km of railways and 2,000 km of waterways) and no natural geographical obstacles such as lake or mountain that may hamper trade between firms. Despite this, we show that productivity distribution is shaped by geography.

The paper proceeds as follows. Section 2 discusses our data sources and presents some empirical evidence that will guide our model presented in Section 3. Section 4 present our empirical design and results. The final section concludes.

---

2. One might be concerned by the fact that Belgium is a very open economy, where import and export amount to 80% of GDP. However, even in a very open economy as Belgium, most of the trade between firms occurs locally. At a more microeconomic level, we also observe that only a small fraction of Belgian firms (around 5% in 2012) are sourcing inputs abroad.
2  Data and descriptive evidence

A full description of the local sourcing strategy of firms is seldomly available. In most countries researchers can access to firm level information on imports by country of origin and by products, but only few information are available regarding domestic transactions between firms. The Compustat data for the US provide some information about the 10 largest suppliers of their sampled firms, and the Japanese data used in Bernard et al. (2017) provide qualitative data on Japanese firms domestic linkages. For this paper, we use a detailed firm level dataset for Belgium. Referred to as the NBB B2B dataset in Dhynne, Magerman and Rubinova (2015), it provides a full description of all domestic transactions between any pair of Belgian VAT affiliates, as long as one of these two affiliates is buying from the second one for at least 250 EUR in any year. This dataset has been used in recent papers (f.i. Dhynne and Rubinova, 2016, Magerman et al., 2016, Bernard et al., 2017, Tintelnot et al., 2017). Observing the amount traded between two firms is a unique feature of the Belgian dataset compared to other datasets (similar data is becoming available for other countries, f.i. Costa Rica). However, we have no information of what is traded between two firms. It can be material inputs, services or capital goods.

The NBB B2B dataset can be viewed as a kind of input-output matrix where each row and each column is a firm. Still, this dataset departs from traditional input output table in many dimensions. First, the way trade intermediaries are recorded in the B2B dataset is fundamentally different from standard IO tables. In standard IO tables and national accounts, the contribution of the wholesalers and retailers to the economy and their intermediate deliveries to other sectors is measured in terms of the value added provided by wholesalers and retailers to the economy. In our transaction data, we observe gross transactions to or from trade intermediaries. The contribution of wholesalers and retailers in the network is therefore much larger than in standard IO tables. Second, as mentioned above, the B2B transaction data does not discriminate between the delivery of inputs (material or services) and of capital goods.

The domestic transaction data can be merged with other datasets (firm level balance sheet data, firm level international trade) that provide firm level information such as its location, its number of establishments, its international trade status, its size and its productivity, and the fact that it may be part of a domestic or international group\textsuperscript{3}. Data on domestic transactions are available from 2002 up to 2014. However, in this paper, we restricted our empirical analysis to one cross-section\textsuperscript{4}.

\textsuperscript{3} Belgian firms are localized according to the ZIP code of their headquarter. Concerning the international trade status, we define three status based on the fact that firm may only import, only export or be involved in both types of foreign trade. The firm size is measured by its employment in full time equivalent and its productivity by the TFP estimated at the NACE Rev.2. 2 digit level using the Wooldridge LP estimator (Wooldridge, 2009). Using data on domestic and international financial linkages, we also observe if a firm belongs to a group or not.

\textsuperscript{4} The results presented are however robusts across various cross-section and a panel analysis limited to a smaller set of firms confirmed our main results.
We first provide some empirical evidence that will guide the development of the model in Section 3. Firms tend to interact mostly with local partners (see Figure 2), with 15% of firm-to-firm relations taking place within a 5 km range. The non-weighted average distance between a domestic supplier and a domestic buyer is 38 km, while this average distance would be 72 km within the Belgian network if relations were randomly chosen.

Figure 1 in Introduction has emphasized the role of the extensive margin of local sourcing. From now on, we restrict our sample to the 311,590 firms for which we know the sector of activity. Although the distribution of the number of suppliers is skewed, firms have on average 29.9 domestic suppliers. Excluding suppliers that are operating in the wholesale, retail, transportation, network industries (electricity, gas and water) or in the construction sector, this average number falls to 11.7. The fixed cost associated with a purchase to the formers, and in particular to the distribution sector, cannot be very high, so we should not expect those sectors to be crucial in explaining the heterogeneity in productivity. Construction is also excluded since it provides capital inputs instead of material inputs. Among the remaining suppliers, about two third are active in support services (NACE Rev 2 55 to 82), and about one third in manufacturing (NACE Rev 2 10 to 33). While domestic sourcing concerns the vast majority of firms, only very few firms are involved in international sourcing. Moreover, it is not a big surprise that manufacturing firms typically have more suppliers and display higher heterogeneity in their sourcing strategy.
Table 1 - Descriptive statistics - Local sourcing strategies in 2012

<table>
<thead>
<tr>
<th>All firms [311,500 firms]</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of suppliers</td>
<td>3</td>
<td>6</td>
<td>15</td>
<td>34</td>
<td>65</td>
<td>29.9</td>
</tr>
<tr>
<td>excluding wholesalers, retailers, network industries, construction and transportation</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>12</td>
<td>25</td>
<td>11.7</td>
</tr>
<tr>
<td>Number of support services suppliers</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>15</td>
<td>7.3</td>
</tr>
<tr>
<td>Number of manufacturing suppliers</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>9</td>
<td>3.6</td>
</tr>
<tr>
<td>Number of source foreign countries</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Manufacturing firms [22,135 firms]</th>
<th>p10</th>
<th>p25</th>
<th>p50</th>
<th>p75</th>
<th>p90</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of suppliers</td>
<td>4</td>
<td>12</td>
<td>32</td>
<td>73</td>
<td>151</td>
<td>63.2</td>
</tr>
<tr>
<td>excluding wholesalers, retailers, network industries, construction and transportation</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>29</td>
<td>64</td>
<td>26.3</td>
</tr>
<tr>
<td>Number of support services suppliers</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>13</td>
<td>29</td>
<td>12.6</td>
</tr>
<tr>
<td>Number of manufacturing suppliers</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>15</td>
<td>32</td>
<td>12.2</td>
</tr>
<tr>
<td>Number of source foreign countries</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Note: Sample size = 311,590 Belgian firms in 2012.
In this paper, we will emphasize the role of services in shaping the productivity distribution. Out of the 311,590 firms in our sample, 292,828 firms outsource at least to one supplier active in the services industries NACE Rev2 55 to 82. Table 2 lists the top ten source services sectors, ranked by the number of buyers. Consistently with the results obtained from the Structure of Enterprises Survey, legal and accounting activities (NACE Rev 2 69) ranks number one based on the number of firms, with 242,593 buyers (82.8% of the sample). For all sectors, however, sector rank based on the number of customers does not equal the rank based on transaction values. This is suggestive of firm level heterogeneity in fixed cost of sourcing to those sectors.

It is useful to assess the extent to which firms follow a hierarchical pecking order in their domestic sourcing behavior. To do so, we follow Eaton et al. (2011) and Antrás et al. (2017) and count the number of firms that source from the top-ranked sector (legal and accounting activities) and no other sector, the number that source from the top two sectors (legal and accounting activities & Office administrative, office support...) and no others, and so on. We calculate these statistics irrespective of firms sourcing outside the top ten sectors. In 2012, 27.5% of firms follow the pecking order listed in Table 2. If we calculate the share of firms that would follow this hierarchy if firms selected into sectors randomly, we would observe a share of 9.3%, which is only one third the share observed in the data (see Table 3). The results on the pecking order are similar to those in Antrás et al. (2017) who showed that 36 percent of US importers follow a pecking order for the top ten destinations, which is higher by a factor 1.8 than what would be predicted under the random entry calculation.

One more result on the pecking order is worth discussing (see Table 3). The compliance with a ranking which is allowed to vary by sourcing sector (27.7%) is not very different from that with the global ranking of services sectors (27.5%) presented in Table 2. This is in contrast to what is obtained if we focus only on manufacturing suppliers. In this case, the compliance with the pecking order is much better if we allow the ranking to be sector-specific (23.8%) than with a global ranking (10.7%). This suggests that, contrary to the purchases to the manufacturing firms, the services sourcing strategy does not depend on the sector of activity, but it is instead common across the economy.

3 Theoretical framework

We develop a model along the lines of Antrás et al. (2017) by exploiting the complementarity mechanism herein. Complementarity between suppliers has the strong implication that there should be a strict hierarchical order in the extensive margin of outsourcing. We however modify their multi-country model to allow for the possibility that firms endogenously chose the set of tasks they will perform in-house. In this model, we assume that, in order to produce a final good, firms need to perform a continuum of tasks that they can decide to do in-house or to outsource. The motivation of outsourcing is production efficiency as firms are heterogeneous in their ability to perform a specific task.
<table>
<thead>
<tr>
<th>Sector of activity</th>
<th>Number of buyers</th>
<th>% of total</th>
<th>Rank by firms</th>
<th>Rank by value</th>
</tr>
</thead>
<tbody>
<tr>
<td>legal and accounting activities</td>
<td>242,593</td>
<td>82.8</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Office administrative, office support and other business support activities</td>
<td>136,838</td>
<td>46.7</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Computer programming, consultancy and related activities</td>
<td>89,428</td>
<td>30.5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Rental and leasing activities</td>
<td>83,486</td>
<td>28.5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Activities auxiliary to financial service and insurance activities</td>
<td>81,239</td>
<td>27.7</td>
<td>5</td>
<td>16</td>
</tr>
<tr>
<td>Activities of head office; management consultancy activities</td>
<td>79,693</td>
<td>27.2</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Financial service activities, except insurance and pension funding</td>
<td>77,098</td>
<td>26.3</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>Services to buildings and landscape activities</td>
<td>75,503</td>
<td>25.8</td>
<td>8</td>
<td>11</td>
</tr>
<tr>
<td>Advertising and market research</td>
<td>67,861</td>
<td>23.2</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>Publishing activities</td>
<td>67,576</td>
<td>23.1</td>
<td>10</td>
<td>14</td>
</tr>
</tbody>
</table>

Note: Top ten source services sectors (NACE Rev 2 55-82)
### Table 3 – Compliance with the peaking order of source sectors

<table>
<thead>
<tr>
<th></th>
<th>Services suppliers</th>
<th>Manufacturing suppliers</th>
<th>All suppliers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Random entry</td>
<td>Data</td>
</tr>
<tr>
<td>All buyers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global ranking</td>
<td>27.5</td>
<td>9.3</td>
<td>17.4</td>
</tr>
<tr>
<td>Buying sector-specific ranking</td>
<td>27.7</td>
<td>tbc</td>
<td>21.2</td>
</tr>
<tr>
<td>Manufacturing buyers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global ranking</td>
<td>29.2</td>
<td>4.3</td>
<td>10.7</td>
</tr>
<tr>
<td>Buying sector-specific ranking</td>
<td>28.3</td>
<td>tbc</td>
<td>23.8</td>
</tr>
</tbody>
</table>

Note: Each cell represents the fraction of firms that follow exactly the peaking order of supplying sectors of activity (2-digits NACE sectors) in the data or in a counter-factual exercise of random selection of supplying sectors of activity.
In our framework, firms will tend to specialize themselves in the set of tasks they are better at performing while they will outsource the remaining tasks. Setting a trading relationship with any supplier implies a fixed cost, which reflects the fact that outsourcing tasks requires some coordination or negotiation with the trading partner. This model of trade in tasks is one of the simplest we can think of that can square with the facts. It enables us to abstract from complex solvability issues, arising in any model of trade in goods due to double marginalization (see Tintelnot et al, 2017). It departs from existing models by dropping the exogenous assumption of upstream and downstream firms. This enables our model to come to terms with some features of the data by allowing firms to sell both at final consumers and at any other firm. Our simplifying framework comes at a cost, however, since in the data we do not observe trade in task but trade in goods and services. Still, we believe that it does not alter the prediction of our model as, when a firm source intermediate inputs from a supplier, this can be viewed as the fact that the firm is sourcing the tasks needed to produce those inputs to the supplier. So for instance, when a car manufacturer sources the tires of its cars to a tire producer, it in fact outsources the task of designing and producing those tires instead of performing those tasks itself. The determinants of the trade relationship between those two firms are still the same than those highlighted in our model.

Market structure

We consider the sourcing strategy of a firm $i$ producing a final good. Suppose firm $i$ owns a blueprint to produce a single differentiated variety of final product. Consumers value the consumption of differentiated varieties of products according to a standard symmetric CES utility function. These preferences give rise to the following demand for firm $i$:

$$q_i = A p_i^{-\sigma}$$

where $A$ is a demand-shifter that the firm treats as exogenous.

Production of final-product varieties require the assembly of a continuum of measure one of tasks, assumed to be imperfectly substitutable with each other, with a constant and symmetric elasticity of substitution equal to $\rho$. The marginal cost of a firm $i$ is

$$c_i = \left( \int_0^1 z_i(t)^{1-\rho} dt \right)^{\frac{1}{1-\rho}}$$

where $z_i(t)$ is the price of an individual task $t$ paid by firm $i$.

5. In 2012, 67% of the firms in our dataset were selling to other domestic firms, 74% to final demand and 53% to both.
Local sourcing

All firms can produce all tasks with labor under constant-returns-to-scale technologies. We denote by $a_i(t)$ the unit labor requirement associated with the production of task $t \in [0, 1]$ by firm $i$. We treat the (infinite-dimensional) vectors of tasks efficiencies $1/a_i(t)$ as the realization of an extreme value distribution. More specifically, the efficiency of firm $i$ in producing a task $t$ is a realization of a random variable from the Frechet distribution $P r(a_i(t) \geq a) = e^{-\psi_i a}$. As in Eaton and Kortum (2002), $\theta$ determines the variability of productivity draws across tasks. These draws are assumed to be independent across firms and tasks. The firm’s core efficiency $\phi_i$ which scales the ability to produce any task is the source of firm-level heterogeneity in our framework. For instance, this parameter may reflect the managerial capabilities of the firm. By assuming that management competence is an important component of firms’ total factor productivity, we follow a growing literature emphasizing the role of management in shaping the patterns of efficiency distribution (Bloom et al., 2012, Bloom et al., 2013, Bloom et al., 2017, Syverson, 2011).

The firm can potentially produce all tasks in-house. Alternatively, the firm can decide to concentrate on its core activities (i.e. the tasks for which it gets the better draws) and outsource the remaining tasks. Trade in tasks however requires the payment of fixed and variable costs. To purchase a bundle of tasks from a particular firm $j$, firm $i$ must incur a fixed cost $f_{ij}$ paid in terms of labor. Furthermore, trade in tasks is subject to iceberg trade costs $\tau_{ij}$.

To simplify matters, we assume that firms sell tasks at their marginal cost. In the Nash bargaining between buyer and supplier, the buyer has therefore the full bargaining power and extracts the entire surplus (Tintelnot et al., 2017). Since by assumption in any firm-to-firm trade the buyer has the full bargaining power, firms do not make profits from sales to other firms. While restrictive, this assumption together with the trade in tasks framework allows us to abstract from complex fixed-point issues.

As a result, the price of an individual task $t$ paid by the firm $i$ is

$$z_i(t) = \min_{j \in J_i} \{a_i(t) \cdot \tau_j a_j(t)\}$$

where $J_i$ is the set of firms for which firm $i$ has paid the associated fixed cost of outsourcing $f_{ij}$.

At this stage, it is worth discussing our framework. The way we introduce trade in tasks is close in spirit to the idea of trade in value-added (Timmer et al., 2014, or Johnson and Noguera, 2012 and 2017). It is motivated by the empirical fact that local production chains look more like spiders than snakes (as labelled by Baldwin and Venables, 2013). The average length of international production chains is close to 3 (see Antràs et al., 2012). As shown in Table 1, in the 2012 Belgian network, the average number of suppliers is 29.9, with 10 percent of

---

6. This fixed cost may for instance represent the cost associated to the set-up of a specific contract between the two firms. In the Belgian context, these fixed cost may also reflect cultural barriers to trade if buyers and sellers are located in different Belgian regions.
firms having more than 65 suppliers. One reason why sequential production is less pervasive than expected is because most of the previous literature has focused on manufacturing processes of production. Services need much less fragmented production chain. Yet, once we consider the whole economy, one should note that trade in service accounts for the largest part of trade between firms. According to IO table for Belgium, trade in service amounts to 70% of the total value.

**Optimal sourcing strategy**

Using the properties of the Frechet distribution, one can show that the firm will source a positive measure of tasks from each supplier in its sourcing strategy set $J_i$. Furthermore, the share of tasks sourced from any supplier $j$ is simply given by

$$\chi_{ij} = \frac{\varphi_j (\tau_{ij})^{-\theta}}{\varphi_i + \Theta_i} \quad (4)$$

Following Antràs et al. (2017), we call $\Theta_i = \sum_{k \in J_i} \varphi_k (\tau_{ik})^{-\theta}$ the sourcing capability of firm $i$ and $J_i$ the sourcing strategy of firm $i$. The overall marginal cost faced by firm $i$ can be expressed as

$$c_i = \gamma (\varphi_i + \Theta_i)^{-1/\theta} \quad (5)$$

where $\gamma = \left[ \Gamma \left( \frac{\theta + 1 - \phi}{\phi - 1} \right) \right]^{1/(\rho - 1)}$ and $\Gamma$ is the gamma function. As we assume that firms act as monopolistic competitors when selling their product to final consumers, they charge a constant mark-up over marginal cost. Using (5), the properties of the Frechet distribution and the constant markup over marginal cost, firms profit can be written as:

$$\pi_i = (\varphi_i + \Theta_i)^{(\sigma - 1)/\theta} B - \sum_j f_{ij} \quad (6)$$

where $B = \frac{\gamma^ {\sigma - 1}}{\sigma - 1} \left( \frac{\sigma}{\phi - 1} \right)^{1 - \sigma} A$. The problem of maximizing (6) is not straightforward to solve because the decision to include a supplier $j$ in the set $J_i$ depends on the number and characteristics of the other suppliers in this set. When deciding whether to add a new supplier $j$ to the set $J_i$, the firm trades off the reduction in costs associated with the inclusion of that supplier in the set $J_i$ against the payment of the additional fixed cost $f_{ij}$. Using a first-order Taylor rule, this net gain of adding supplier $j$ can be approximated by:

$$\Delta_{ij} (\Theta_i) = \frac{\sigma - 1}{\theta} (\varphi_i + \Theta_i)^{\sigma - 1 - \theta} \varphi_j (\tau_{ij})^{-\theta} B - f_{ij} \quad (7)$$

The marginal gain from adding a supplier is higher for more productive firm whenever $\sigma - 1 > \theta$, which we henceforth assume. In this case, efficiency and outsourcing display complementarities, which is consistent with the Figure 1 in
Introduction. If, instead, $\sigma - 1 < \theta$, more efficient firms outsource less tasks as they find more profitable to capitalize on their comparative advantage in producing tasks more efficiently.

Interestingly, under the condition that $\sigma - 1 > \theta$, Equation (7) also exhibits complementarities between suppliers. A firm with a larger sourcing strategy will find more profitable to outsource tasks to one more supplier. This is more likely when consumer demand is elastic and efficiency draws are heterogeneous. This feature of the model was already in Antràs et al. (2017) and explains why firms would follow a predicted pecking order. Firms rank their potential suppliers according to their attractiveness, which is simply dictated by $\varphi_j (\tau_{ij})^{-\varphi}$ and $f_{ij}$. The pecking order in terms of sectors, and in particular in services sectors, suggests that suppliers within the top ranked sector tend to exhibit favourable productivity associated with relatively low fixed cost. They are chosen in priority by firms that seek to outsource tasks. There are however reasons why this hierarchy is not strictly followed by firms. Two perfectly similar firms would strictly follow a predicted pecking order. But firms differ in various respects. They have their own specific draws of tasks efficiencies. They also differ in terms of geography (trade cost) and relation-specific fixed costs.

Our model also puts us in a position to derive some testable predictions. From Equations (7) and (4), we have that:

More efficient and less distant suppliers are more likely to be chosen. The share of tasks is increasing in the supplier’s productivity and decreasing both in the geographic distance between the supplier and the buyer and in the buyer’s productivity.

4 Econometric results

4.1 The determinants of the decision of local sourcing

To confront our model with data, we use the NBB B2B dataset described in section 2. As mentioned above, this dataset provides the researcher with a complete description of all the business relationships managed by Belgian firms. For every registered corporation in Belgium, we therefore have a complete view of its domestic suppliers. Simultaneously, we know which Belgian firms were not considered by a firm as a potential supplying partner.

Based on our model, the selection of a specific trading partner relies on both geographical (i.e. the distance between the firm and a potential supplier) and economical (the level of economic performance of both trading partners) factors. In this section, we rely on the estimation of a Probit equation characterizing the decision of firm $i$ to source inputs from firm $j$ to test the empirical predictions of our model.

In our model, firms that trade domestically may face some trade costs. Geography, as illustrated in Figure 2, shapes the structure of the Belgian production networks. Culture may also be an important determinant of the organization of the domestic production network. Belgium is indeed a country with 3 regions,
with Flanders being Dutch-speaking, Wallonia French-speaking and Brussels bilingual. Therefore, even if there is no tariff barriers to trade within Belgium, there can be sizeable fixed costs of trade between firms located in different regions (Flanders versus Wallonia).

Therefore, following our modeling strategy and Equation (7), we assume that the probability that \( j \) belongs to the set of suppliers of firm \( i \) is given by

\[
P[I (Sales_{ijt} > 0) | X_{ijt}] = \Phi (\beta_0 + \beta_1 \text{Dist}_{ij} + \beta_2 \# \text{Lang}_{ij} + \beta_3 \text{TFP}_{i0} + \\
\beta_4 \text{TFP}_{jt} + \beta_5 L_{i0} + \beta_6 L_{jt} + \beta_7 \text{Participation}_{ijt} + ...)\
\]

where:
- \( \text{Dist}_{ij} \) is the distance between the Belgian headquarters of firm \( i \) and its potential supplier \( j \)'s Belgian headquarters;
- \( \# \text{Lang}_{ij} \) is a binary variable reflecting the fact the Belgian headquarters of firms \( i \) and \( j \) are located in two regions not sharing a common language (either Flanders or Wallonia);
- \( \text{TFP}_{i0} \) is the total factor productivity of firm \( i \) measured in a pre-sample period, in order to reduce the potential influence of the sourcing strategy decision in time \( t \) with the level of efficiency of the buyer. This measure should reflect the \( \phi_i \) parameter as the ex-ante ability of the firm to combine all the tasks needed for its production;
- \( \text{TFP}_{jt} \) is the logarithm of the current total factor productivity of firm \( j \), as the current decision of firm \( i \) to source from firm \( j \) should depend from its current level of efficiency;
- \( L_{i0} \) and \( L_{jt} \) are respectively the employment of firm \( i \) in a pre-sample period or of the current level of employment of firm \( j \);
- \( \text{Participation}_{ijt} \) is a binary variable reflecting the fact that there exist a financial participation between firm \( i \) and the potential supplier \( j \) at time \( t \);
- Additional controls includes the internation trade status of both firms, dummies reflecting the sector of activity of both firms, dummies characterizing the zip code of both firms,...

In order to estimate a Probit, one needs both “0”s and “1”s. If our transaction dataset provides us with all the “1”s in a given year, we need to sample the “0”s. Indeed, we cannot put all the “0” in our sample when estimating the Probit equation. Considering only the 2012 cross-section and the subset of 110,129 firms for which we observe TFP, this would involve analyzing the 5,023,543 observed transactions and 12,123,362,969 pairs of buyers and potential sellers that did not trade together. Such a large amount of data cannot be handled on standard computers and prohibits the estimation of a Probit equation using the full set of observed and potential transactions. To perform our estimation, we randomly selected a given number of potential transactions (effective or not) for any Belgian firm in our dataset. We have built three samples for the estimation of our baseline regression.

The first one considers either as a buyer or as a potential supplier all the firms included in our sample for which we observe the location of the headquarters,
Table 4 – Sourcing choice of firm $i$: dependent variable: $I(sales_{ij,t} > 0)$

<table>
<thead>
<tr>
<th></th>
<th>All suppliers</th>
<th>Manufacturing suppliers only</th>
<th>Service suppliers only</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\ln \text{Distance}_{ij}$</td>
<td>-0.242***</td>
<td>-0.964***</td>
<td>-0.264***</td>
</tr>
<tr>
<td>($\hat{0.0007}$)</td>
<td>($\hat{0.0017}$)</td>
<td>($\hat{0.0010}$)</td>
<td>($\hat{0.0015}$)</td>
</tr>
<tr>
<td>$\neq \text{Lang}_{ij}$</td>
<td>-0.247***</td>
<td>-0.627***</td>
<td>-0.236***</td>
</tr>
<tr>
<td>($\hat{0.0016}$)</td>
<td>($\hat{0.0200}$)</td>
<td>($\hat{0.0032}$)</td>
<td>($\hat{0.0125}$)</td>
</tr>
<tr>
<td>$\ln \text{TFP}_{0}$</td>
<td>0.054***</td>
<td>0.216***</td>
<td>0.258***</td>
</tr>
<tr>
<td>($\hat{0.0077}$)</td>
<td>($\hat{0.0233}$)</td>
<td>($\hat{0.0032}$)</td>
<td>($\hat{0.0125}$)</td>
</tr>
<tr>
<td>$\ln \text{TFP}_{i,t}$</td>
<td>0.076***</td>
<td>0.302***</td>
<td>0.061***</td>
</tr>
<tr>
<td>($\hat{0.0099}$)</td>
<td>($\hat{0.0232}$)</td>
<td>($\hat{0.0032}$)</td>
<td>($\hat{0.0125}$)</td>
</tr>
<tr>
<td>$\ln L_{i0}$</td>
<td>0.087***</td>
<td>0.347***</td>
<td>0.110***</td>
</tr>
<tr>
<td>($\hat{0.0099}$)</td>
<td>($\hat{0.0232}$)</td>
<td>($\hat{0.0032}$)</td>
<td>($\hat{0.0125}$)</td>
</tr>
<tr>
<td>$\ln L_{ij,t}$</td>
<td>0.127***</td>
<td>0.505***</td>
<td>0.092***</td>
</tr>
<tr>
<td>($\hat{0.0043}$)</td>
<td>($\hat{0.0233}$)</td>
<td>($\hat{0.0032}$)</td>
<td>($\hat{0.0125}$)</td>
</tr>
<tr>
<td>$\text{Participation}_{ij,t}$</td>
<td>1.489***</td>
<td>157.7***</td>
<td>2.100***</td>
</tr>
<tr>
<td>($\hat{0.114}$)</td>
<td>($\hat{0.0238}$)</td>
<td>($\hat{0.0238}$)</td>
<td>($\hat{0.171}$)</td>
</tr>
</tbody>
</table>

$i$ and $j$ zip dummies  | Yes          | Yes                          | Yes                    |
$i$ and $j$ sector dummies | Yes         | Yes                          | Yes                    |
Additonal controls       | Yes          | Yes                          | Yes                    |
Observations             | 11,593,878   | 4,823,339                    | 12,154,809            |

Note: Based on the 2012 cross-section, $I(sales_{ij,t} > 0)$ is a binary variable that indicates whether firm $i$ sources inputs from firm $j$ at time $t$. $\text{Distance}_{ij}$ is the "as the crow flies" distance in km. $\neq \text{Lang}_{ij}$ is a binary variable indicating that firms $i$ and $j$ do not share a common language. $\text{TFP}$ is total factor productivity (estimated at the NACE 2-digit level using the Woodford-IP estimator). Year 0 refers to the first year of observation in the data. $L$ is the number of employees, in FTE. $\text{Participation}_{ij,t}$ is a dummy variable indicating that $i$ or $j$ owns at least 50% of the capital of the other firm. Additional controls include sectoral dummies for $i$ and $j$, dummies for the degree of internationalization of $i$ or $j$ (exporter, importer, MNE) or if $i$ or $j$ are multilant firms. Standard errors of the estimated coefficients are clustered at the sourcing firm level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 


its employment and an estimate of its total factor productivity, either in its first year of observation in our full dataset\textsuperscript{7} or in 2012. This sample covers all sectors of activity, from manufacturing to services including wholesalers and retailers and network industries. For this sample, we consider 100 randomly selected potential suppliers for every firm, ending up with 11,593,878 potential supplying relations, out of which 4,619,255 effectively ongoing.

The second one only considers the subset of manufacturing firms as suppliers and firms that source from this subset of suppliers as buyers. Because the subset of suppliers and buyers is smaller, we select 500 potential transactions for each buyer. This sample covers 4,823,339 potential transactions, out of which 678,427 are real transactions.

Finally, a third sample only considers the subset of firms active in service sectors NACE 55 to 82 as suppliers and firms that source from this subset as buyers. We also conserved 500 potential transactions for each buyer. This sample covers 12,154,809 potential transactions with 1,038,891 real transactions.

\textsuperscript{1}'s are naturally over-represented in our three samples, the true probability of a transaction being equal respectively to 0.000394 in the first sample, 0.000561 in the second one and 0.000325 in the last one. Therefore, we estimated a weighted Probit equations that corrects for this feature of our datasets. Results are summarized in Table 4. In addition to the estimated coefficients, we also present the estimated elasticity of the probability of a transaction with respect to our set of explanatory variables to allow for an easier comparison of the results.

The results obtained are basically in line with the predictions of our model. Most productive firms tend to source from more suppliers, and better suppliers are more likely to be selected. Distance is also a key determinant of the likelihood of a business relationship. Our results are also consistent with the fact that trade costs may embodied cultural barriers to trade as a common language increases the probability to trade.

Comparing results across sectors, we observe that a supply relationship between two firms located on either side of the linguistic border is particularly difficult when services are exchanged instead of manufacturing goods. This reflects the fact that personal contacts are a true component of a service transaction while trading goods do not necessarily requires a common language. Considering the effect of distance, the impact of distance is also smaller for the decision to supply manufacturing goods and services.

Finally, vertical integration is naturally a strong determinant of business transaction, implying almost always a transaction between the parent and the affiliate\textsuperscript{8}.

\textsuperscript{7} We have firm level data that allows us to estimate a production function and a firm-level TFP for the 2002-2012 period using the Wooldridge method.

\textsuperscript{8} This result raises the issue of the definition of the firm’s border in the empirical literature. The increased phenomenon of spin-offs and the organisation of some firms among multiple legal units for any kind of fiscal or organization reason, especially of large corporations challenges the way we not only define business transactions (is it a relevant sourcing decision ?) but, in other contexts, how we measure firms performance. Ideally, a researcher would like to observe the flows between production units and the characteristics of those production units
These results are robust to alternative specifications involving different level of localisation dummies (at the district level) or firm fixed effects. To check the robustness of our results, and especially of the results related to our geographical variables, we’ve estimated a logit equation instead of a probit equation with buyer and seller specific fixed effects. Because of the large number of coefficients evolved in those specification, the two dimensional fixed effect logit specification is preferred as it can be efficiently estimated using a linear transformation with `reglfe`. These estimations have been conducted for each two digit sector separately on the supplier side and only include transaction \( ij \) specific explanatory variables (\( Distance_{ij} \) and \( \neq Lang_{ij} \)). This allows us to include all the “0” in our specification for each supplying sector. In this case, we model carefully why firm \( i \) has decided to source some tasks from firm \( j \) and not from firm \( k \), while \( k \) and \( j \) are in the same industry. The results related to the impact of distance and of the cultural barriers for manufacturing and service industries are summarized in Figure 3. Globally, they are in line with our baseline results in Table 4.

Figure 3 – Elasticity with respect to \( Distance_{ij} \) and \( \neq Lang_{ij} \) based on FE Logit estimates by supplying sector

From this sectoral analysis, it seems that the impact of distance, even if it is always sizeable, varies strongly across sectors both within the manufacturing or the services sectors but do not differ significantly across those two broad sectors. However, the influence of the difference in the mother language seems to be larger for services. If the semi-elasticity associated to that variable seems to be

\[ \text{employment, production} \]

but in most case, what is available is information at the level of a legal unit (in Belgium, a VAT unit) and according to the organisational form of the firm, you may end up either observing transactions between multiple VAT numbers or no intra-firm trade.
Table 5 - Contribution of firm \( j \) to firm \( i \)'s total sales: dependent variable: 
\[ \ln \frac{\text{sales}_{ijt}}{\text{sales}_{it}} \]

<table>
<thead>
<tr>
<th></th>
<th>All suppliers</th>
<th>Manufacturing suppliers only</th>
<th>Services suppliers only</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ln \text{Distance}_{ij} )</td>
<td>-1.669***</td>
<td>-1.856***</td>
<td>-1.776***</td>
</tr>
<tr>
<td>( \neq \text{Lang}_{ij} )</td>
<td>-1.722***</td>
<td>-1.674***</td>
<td>-2.244***</td>
</tr>
<tr>
<td>( \ln \text{TFP}_{it} )</td>
<td>-0.144***</td>
<td>-0.130***</td>
<td>-0.106***</td>
</tr>
<tr>
<td>( \ln \text{TFP}_{jt} )</td>
<td>0.604***</td>
<td>0.489***</td>
<td>0.534***</td>
</tr>
<tr>
<td>( \ln L_{i0} )</td>
<td>0.070***</td>
<td>0.297***</td>
<td>0.138***</td>
</tr>
<tr>
<td>( L_{jt} )</td>
<td>0.784***</td>
<td>0.587***</td>
<td>0.845***</td>
</tr>
<tr>
<td>( \text{Participation}_{ijt} )</td>
<td>11.270***</td>
<td>15.268***</td>
<td>15.186***</td>
</tr>
</tbody>
</table>

\( i \) and \( j \) Zip code dummies: Yes
Year dummies: Yes
Additional controls: Yes
Observations: 11,593,795

Note: Based on the 2012 cross-section. \( \text{sales}_{ijt} \) is the amount in EUR of tasks / inputs sourced by firm \( i \) from firm \( j \) at time \( t \). \( \text{Distance}_{ij} \) is the "as the crow fly" distance in km. \( \neq \text{Lang}_{ij} \) is a binary variable indicating that firms \( i \) and \( j \) do not share a common language. \( \text{TFP} \) is total factor productivity (estimated at the NACE 2-digit level using the Woodridge-LP estimator). Year 0 refers to the first year of observation in the data. \( L \) is the number of employees, in FTE. \( \text{Participation}_{ijt} \) is a dummy variable indicating that \( i \) or \( j \) owns at least 50% of the capital of the other firm. Additional controls include sectoral dummies for \( i \) and \( j \), dummies for the degree of internationalization of \( i \) or \( j \) (exporter, importer, MNE) or if \( i \) or \( j \) are multiplant firms. Standard errors of the estimated coefficients are clustered at the sourcing firm level. Significance levels: *** \( p < 0.01 \), ** \( p < 0.05 \), * \( p < 0.1 \).

below 0.6 in manufacturing, in most service sectors, it is above 0.7.

4.2 The intensive margin of local sourcing

In section 4.1, we investigate the determinants of the decision to establish a business relationship with a potential supplier. In this section, we exploit the value of the transaction to analyse how firms and geographic characteristics affect the amount supplied by firm \( j \) to firm \( i \). We therefore estimate a Eaton-Kortum Tobit equation of the amount traded between \( i \) and \( j \), censoring our dependent variable to the minimum amount delivered by supplier \( j \) to any firm in the Belgian network. The dependent variable is expressed in relative term with respect to the total sales of \( i \). We then take the log of that share. Results are presented in Table 5.

The results obtained are in line with Equation (4) of our model. When firms source from remote locations, they tend to trade less because of increasing variable costs. Similarly, variable costs are associated to the cultural barriers. Controlling for their total sales, firms naturally source more tasks to more efficient or bigger suppliers. The degree of "intrinsic efficiency" of the firm, proxied by the pre-sample TFP of the buyer, however reduces the relative amount traded with the other firms. This reflects two phenomenons. First, as more efficient
firms source from more suppliers, each of them may represent a smaller share of the total sales of firm $i$. Second, as $i$ is more intrinsically efficient, it’s own contribution to the value of the final good may be larger. In terms of sectoral differences, it seems that transportation costs are larger for the supply of manufacturing inputs while cultural barriers are more effective when services are exchanged. The large impact of vertical integration on the amount of trade also strengthens the argument of the relevance of the definition of the boundary of the firms when analysing B2B transactions.

5 The geography of TFP heterogeneity

As shown in section 4, the productive efficiency of both the firms and their potential suppliers significantly affect their domestic sourcing strategy. In return, the selected sourcing strategy affects the efficiency of the firm as illustrated in Figure 1.

However, as geography also affects the sourcing strategy of the firm, the ex-post production efficiency of a firm may be related to its geographical location in the production network. Firms located in the center of the domestic production network (for instance in Brussels) may enjoy easier access to potential suppliers as the average distance with any Belgian firms may be relatively small compared to firms located in the Ardens, near the French and German borders, or at the seaside. Firms located in the bilingual Region of Brussels are also immune to any cultural barriers. Because of that easier access, the post-sourcing level of TFP of firms located in Brussels may be significantly larger than those of a similar firms located in Bastogne or De Pann.

To test for this assumption, we used our baseline regressions to compute a measure of connectiveness at the zip code level, that would reflect how geography affects the probability of a given firm to connect to any firm included in our sample, neutralizing for the firm specific characteristics. For a given zip code $k$, we compute:

$$P_k = \sum_j \left( \hat{\beta}_1 \text{Dist}_{kj} + \hat{\beta}_2 \neq \text{Lang}_{kj} \right)$$

where:
- $\text{Dist}_{kj}$ is the distance between zip code $k$ and supplier $j$;
- $\neq \text{Lang}_{kj}$ is a binary variable reflecting the fact the firms located in zip code $k$ and firm $j$ are located in two regions not sharing a common language (either Flanders or Wallonia).

We compute this connectiveness considering either the access to any kind of suppliers or to service suppliers only and we rescale our results so our measure ends up between 0 (the least connected zip) and 1 (the no trade frictions case). Both measures are highly correlated indicating that the provision of supporting activities to firms do not seem to suffer from a different geographical distribution than the supply of others inputs. Such a measure naturally only consider the domestic network, which may not be the relevant sourcing area for firms located close to the Belgian borders. For firms located in Bastogne, the relevant sourcing
Figure 4 – Domestic and imported connectiveness at the zip level

area might be extended to part of Germany or France or Luxembourg. If this is indeed the case, our measure of connectiveness overestimate the difficulty for a firm located in non central area to connect with potential suppliers.

This measure of connectiveness can be viewed as a production externality parameter as in Redding and Rossi-Hansberg (2017). The larger it is the more a firm located in zip code k is potentially exposed to the other Belgian firms. To test this assumption, we include this measure in a TFP equation at the firm level. We observe that both global and service connectiveness are positively correlated with the TFP of a firm indicating that firms located in zip codes easily connected to firms in general or to service suppliers in particular tend to experiment higher levels of productivity.

Based on our estimates, firms located in the most connected area are 14.3 or 14.8% more efficient than firms located in the least productive area. Removing totally trade frictions between domestic firms would increase the productivity of the least connected area by 29.6 or 34.0%. On average, Belgian firms would gain 6% of efficiency by reducing trade frictions to the level of the best connected area. It is worth mentioning that if non-central areas were able to overcome their bad domestic connectiveness through imports, firstly we would observe a large fraction of importers in non central areas, which does not seem to be the case (see Figure 4), secondly this would play against our results as these low connected areas would tend to have more productive firms than expected. It is also interesting to highlight the productivity losses related to the cultural barriers to domestic trade. Based on our measure of cultural barrier, firms established in the Flemish and the Walloon regions would already significantly benefit from the disappearance of any cultural barriers to trade. They would record productivity gains of 1.5% in Flanders and 4.0% in Wallonia.

We also considered the sectoral heterogeneity in the trade frictions described in Figure 3. Considering sector specific estimates for the distance and the cul-
<table>
<thead>
<tr>
<th></th>
<th>Using common distance and cultural border effects</th>
<th>Using sector specific distance and cultural border effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{i\in i}$</td>
<td>(1) 0.296***</td>
<td>(3) 0.360***</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.073)</td>
</tr>
<tr>
<td>$P_{i\in j}$</td>
<td>(2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0.340***</td>
</tr>
<tr>
<td></td>
<td>(0.074)</td>
<td>(0.074)</td>
</tr>
<tr>
<td>$\ln L_i$</td>
<td>(3) 0.154***</td>
<td>(4) 0.154***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>113,009</td>
<td></td>
</tr>
<tr>
<td>Sector dummies</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Additional controls</td>
<td>Yes</td>
<td></td>
</tr>
</tbody>
</table>

Note: Based on the 2012 cross-section. Additional controls include sectoral dummies for $i$ and $j$, dummies for the degree of internationalization of $i$ or $j$ (exporter, importer, MNE) or if $i$ or $j$ are multinationals. Standard errors of the estimated coefficients are clustered at the zip level. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. 

Table 6 – TFP and connectivity: dependent variable: $\ln TFP_{it}$
tural border coefficients, we obtain that firms located in the least connected area suffer from a productivity handicap of 15.5% and would be around 35% more efficient in the absence of geographical domestic trade frictions. Totally removing cultural barriers in manufacturing and business services only would generate an increase in TFP of 1.7% in firms located in Flanders and 5.1% in Wallonia.

6 Conclusion

In this paper, we documented the local sourcing strategy of Belgian firms. Exploiting an exhaustive dataset on firms' buyer-seller linkages in Belgium, we are able to fully describe the sourcing behaviour of each individual firm. First, we provide descriptive statistics on the distance between the supplier and the buyer by stressing the fact that most firms tend to trade with local partners. Perhaps more surprisingly, selection into local sourcing exhibits complementarities across sectors of activities. As a result, firms follow a predictable pecking order in their sourcing decision. Even though most firms buys inputs from suppliers operating in the wholesale, retail, transportation or network industries (energy provider or telecommunications), inputs from services sectors and, to a lesser extent, from the manufacturing sectors are crucial in explaining the heterogeneity in the extensive margins of local sourcing.

Based on a modified version of the global sourcing model of Antras et al. (2017), we model local sourcing as trade in tasks, with firms deciding on the set of tasks they want to keep in house and the ones they want to outsource to other (domestic) firms. Under this framework, more efficient and less distant suppliers are more likely to be chosen. The share of tasks is increasing in the supplier's productivity and decreasing both in the geographic distance between the supplier and the buyer and in the buyer's productivity.

The results obtained in the empirical part are basically in line with the predictions of the model. These results are robust to alternative specifications involving different level of localisation dummies or firm fixed effects. From a sectoral analysis, we show that the impact of distance, even if it is always sizeable, varies strongly across sectors both within the manufacturing or the services sectors but do not differ significantly across those two broad sectors.

We finally show that the location of the firm has some strong impact on this performance as it affects its connectivity in the Belgian network. Firms located in places where connectivity to suppliers is high exhibits higher productivity levels. Our estimates suggest that total factor productivity in Belgium would be on average 5.5% higher if all firms were as well connected as in the best connected area in Belgium.
References


