Export Entry and Network Interactions

Evidence from the Belgian Production Network

*Emmanuel Dhyne[◊], Philipp Ludwig[△] and Hylke Vandenbussche[§]

February 2022

Preliminary and incomplete - please do not cite or circulate

Abstract

Low export participation of firms across countries is typically related to high entry costs allowing only the most productive firms to serve foreign markets. In this paper, we move beyond individual firm characteristics to explain export participation and investigate whether firms' domestic network linkages can facilitate export entry. Firms receive information from business interactions with experienced exporters which lowers sunk entry costs and allows them to enter the foreign market. Using rich data of buyer-seller linkages in the Belgian production network, we find that network heterogeneity is a key determinant of the extensive margin of trade. Each additional export signal received via network linkages increases the entry probability by 0.5 – 1.4 percentage points, giving firms with suitable networks a key advantage in accessing foreign markets. The marginal impact of network externalities decreases in network size which we attribute to negative assortative matching in the underlying network formation process.

Keywords: Export entry, buyer-seller network, information frictions, trade barriers, heterogeneous firms

JEL Classification: F12, F13, L14, D83

---

*We thank Ronald Davies, Keith Head, Christian Proebsting, Gonzague Vannoorenberghe, Jaume Ventura and participants of the CES-LCIS-UCD seminar on trade and taxation and the Leuven-Louvain trade workshop for valuable comments and discussions. Philipp Ludwig gratefully acknowledges funding from the Research Foundation Flanders (FWO) thanks the NBB for the hospitality during a research stay where this paper developed.

[◊]National Bank of Belgium and University of Mons. Emmanuel.dhyne@nbb.be

[△]Department of Economics, KU Leuven. philipp.ludwig@kuleuven.be

[§]Department of Economics, KU Leuven. hylke.vandenbussche@kuleuven.be

1 Introduction

Despite continued trade liberalization export participation remains an exclusive phenomenon with only 4-5% of firms directly engaged in cross-border trade (Bernard et al., 2007; Dhyne et al., 2015). This concentration of economic activity at the extensive margin of trade has far-reaching consequences. Low export participation not only weakens competition in domestic markets by allowing a small number of exporters to consolidate market power (De Loecker and Warzynski, 2012) but also severely restricts aggregate export growth which crucially depends on the entry of new firms (Eaton et al., 2009a). The key question is therefore to understand why only a handful of firms can overcome entry barriers and access foreign markets.

This paper aims to contribute to this debate by investigating which factors ultimately determine a firm's entry decision. Traditionally, low export participation has been related to the presence of sunk entry costs allowing only a select number of firms to enter and remain profitable in a foreign market. This insight sparked an influential literature of heterogeneous firm trade models which directly links export entry decisions to individual firm characteristics. Prominent attributes determining export participation are firm productivity (Melitz, 2003), access to finance (Manova, 2013) or previous experience in similar markets (Morales et al., 2019). What all these examples share is that they relate export participation to firm-level characteristics. Firms however do not operate in isolation but constantly interact with other firms in their production network. If information diffuses along network linkages, this interaction exposes firms to the export experience of their peers. As a consequence, accessing foreign markets becomes easier as entry-related sunk costs are partially offset by the received network information. In this setting, each firm's individual network serves as a conduit of export-related information creating additional heterogeneity in entry behavior beyond firm-level characteristics.

Our paper allows for this network-dimension of firm heterogeneity and empirically investigates whether export participation is facilitated by buyer-seller interactions in domestic production networks. Using detailed data from the universe of Belgian firms, we capture each firm's entire domestic production network and export behavior for the years 2002 − 2014. To trace the diffusion of export-related information along buyer-seller linkages, we combine export and network data to detect changes in firms' export experience. Each time a firm starts to export to a new destination, it gathers entry-related information which diffuses through business interactions to the firm's local production network. At the receiving end, these export market signals lower market access costs and increase the probability of the receiver to start exporting to the foreign market in the following year. Our approach therefore exploits the export behavior of peers in a firm's network to detect export market signals and investigates whether incoming signals increase the entry probability to the same market.

To formalize this idea, we introduce network interactions to a stylized entry model in which firms are heterogeneous along both a firm-dimension, capturing individual characteristics, and a network-dimension, capturing their linkages to other firms. Export decisions are based on a simple trade-off between the costs and benefits of entering a foreign market. The key novelty is to express sunk entry costs as a function of network linkages using functional form assumptions inspired by the literature of social networks (Bramoullé et al., 2009) and spatial economics (Anselin et al., 2008). As a consequence, differences in network linkages directly translate into different entry costs creating a second dimension of firm heterogeneity beyond firm productivity.

To achieve this, we treat network effects as an externalitiy which implies that network formation is not strategic and conditionally exogenous. This assumption is relaxed at a later stage. While restrictive, the stylized setting allows us to study both dimensions of heterogeneity at the same time and empirically investigate the relevance of network interactions for export participation.
The resulting estimation equation takes the form of a time-space recursive model which relates export entry decisions to individual firm characteristics and incoming export signals. Since networks not only contribute to the diffusion of valuable export signals but also expose firms to unrelated information we make two important adjustments. First, we distinguish between matching and non-matching signals based on whether the signal origin coincides with the destination the export entry occurs in. Only matching signals are expected to lower market access costs and facilitate entry, meaning an export start to China should not benefit from signals originating in the US market. Second, we consider two separate model specifications which differ in their treatment of network noise. While signal intensity ignores any negative impact of noise for network effects, signal clarity discounts received export information by the amount of noise in the network.
In both cases, a precise estimation of network effects relies on a combination of high-dimensional fixed effects and network instruments to ensure that any correlation in firm responses is caused by network interactions and not driven by common unobserved shocks. While network linkages are initially considered exogenous, we account for endogenous network formation in an extension by introducing a network selection model as in Arduini et al. (2015) and Qu et al. (2017) to our benchmark setting. This allows us to correct for a potential bias in estimated network effects from correlated linkage and entry decisions. Using detailed balance sheet, trade transaction and network data from the universe of Belgian firms, we then estimate our entry model for export starts occuring between 2004-2014.
Our preliminary results indicate that network externalities play a decisive role in export entry decisions of Belgian firms. First, each additional matching signal increases the entry probability between 0.5 – 1.4 percentage points while non-matching signals do not have any effect. As firms can receive multiple matching signals in any given period, the economic impact of this channel can be considerable and demonstrates that network heterogeneity is an important new dimension for the extensive margin of trade. These benchmark estimates are robust to sample selection and remain unchanged across competing estimation methods. Second, we find that the marginal effect of network externalities decreases significantly in firm size. While all firms benefit from additional export signals, the impact is 3-4 times stronger for small compared to large firms. We show that these differences can be traced back to negative assortative matching in the underlying network formation process which exposes large firms to a disproportionate amount of network noise. Hence, despite receiving more export signals due to their large network, the marginal impact of each signal appears to be attenuated by the high degree of network interference. These diseconomies of network scale stand in stark contrast to firm-level characteristics like productivity which show constant marginal effects across subsamples of small and large firms. This suggests that network effects might be a particularly important tool to facilitate a stronger participation of small firms in foreign markets.
While our findings emphasize that network interactions are an important factor in understanding firm behavior at the extensive margin of trade, networks and entry decisions have typically been considered as separate phenomena in the international trade literature. Studies investigating the

determinants of export participation mostly focus on firm-level characteristics. Traditionally, firm productivity and firm size have received most attention to explain the sorting of firms into exporters and non-exporters in accordance with the canonical heterogeneous firm trade model of Melitz (2003). Recent contributions also highlight the role of experience generated from a firm's past export activity. This helps to account for correlated export patterns across time (Albornoz et al., 2012) and space (Morales et al., 2019) as entering foreign markets occurs more frequently if firms are already active in the region. Similarly, Arkolakis et al. (2021) find that firm experience can occur in form of scope economies at the product level as market-access costs fall with the number of products the firm already exports to the destination. We directly contribute to this literature by highlighting that expertise generated from past export activity not only benefits firms' own export expansion but also diffuses through network linkages and thereby becomes an important determinant for export participation of others. Other insights from previous studies are directly mirrored in our estimated network effects. Non-matching export signals for example still have a mildly postive impact on export entry if they originate in markets with close geographic proximity to the actual entry destination.

A separate literature in international trade has started to formally integrate buyer-seller linkages into existing trade models to provide a theoretical foundation for observed firm behavior at the extensive margin of trade. Chaney (2014) explains spatially correlated entry patterns of firms with a search mechanism that allows firms to use their existing contacts in foreign markets to expand to nearby destinations. A further export expansion of the firm is therefore facilitated by existing cross-border linkages. Eaton et al. (2016) and Bernard et al. (2018) look at the microfoundation of the extensive margin of trade by directly modeling the formation of international buyer-seller linkages. In contrast to these papers, we study domestic buyer-seller linkages and allow them to facilitate the creation of cross-border linkages in form of export starts. As export starts are only captured at the firm-destination level, we abstract from the identity of the importing firm in the foreign market. Nevertheless, we contribute to the literature on international buyer-seller linkages in an important way. By highlighting that diffusion in domestic networks facilitates export entry, we show that domestic networks actively contribute to the formation of international networks.

Closest to this paper is a third strand of empirical literature which directly relates export decisions to activity in domestic networks. A clear point of distinction among papers in this group is the way authors try to capture domestic firm-to-firm interactions. A common approach in the export spillover literature (Fernandes and Tang, 2014; Koenig, 2009) is to extrapolate unobserved network linkages from spatial proximity which implicitly assumes that interactions are limited to the specified geographic unit and occur between all firms within it. Our production network approach in contrast is based on observed buyer-seller transactions which mitigates concerns of network misspecification common to geographic proxies. Recent alternatives have been presented by Patault and Lenoir (2021) who study if hiring employees from exporters increases the likelihood to start serving the same foreign clients and by Connell et al. (2019) who explore if indirect exporting through wholesalers facilitates subsequent export entry. While human capital movements and learning through wholesalers are important mechanisms to access export-related information, they are far more restrictive than our approach which allows information to diffuse in absence of labor movements and through any business interaction.

Finally, this paper is related to a vast literature on peer effects in networks (Advani and Malde, 2018; Bramoullé et al., 2020). Depending on the setting, the relationship between agent and network behavior is either characterized by homophily, meaning agents seek to conform to average network behavior or strategic complementarity, where agent and network behavior is mutually reinforcing. Introducing these ideas to study export participation leads to two important methodological contributions. First, we show that in an information diffusion setting, choosing between strategic complementarity and homophily translates to ignoring or accounting for network noise. Second, this distinction is particularly important if interactions are based on production networks as negative assortative matching leads to diverging network effects under different behavioral assumptions. In case theory cannot determine a priori which option is more appropriate, researchers need to be mindful of the underlying network formation process when deciding how to model agent behavior.

This paper has 7 sections. Section 2 presents a stylized model of export entry, shows how to introduce network interactions and discusses strengths and weaknesses of the approach. Section 3 describes our data sources and sample selection. Section 4 discusses identification and estimation of our augmented entry model. Our preliminary results are presented in section 5 and discussed in section 6. The last section concludes.

2 Theoretical framework

Our goal is to assess whether foreign market access not only depends on each firm's individual characteristics but can also be facilitated by information externalities originating from a firm's domestic production network. We therefore need a theoretical framework which relates export entry decisions to two dimensions of firm heterogeneity: firm productivity and network interactions. While the former is a standard component in most trade models since the seminal contribution of Melitz (2003), heterogeneity in firm networks only recently attracted attention in the trade literature[1] and has commonly not been considered as a direct determinant of export participation[2].

We therefore proceed in two steps. First, we use the model of Koenig (2009) to illustrate the key determinants of firm export entry behavior in a standard setting which only allows firms to differ in productivity. More productive firms can expect higher profits in foreign markets and are therefore more likely to enter. In a second step, we then introduce network interactions to the model by allowing entry barriers in foreign markets to change in response to export-related information received from network peers. In this setting, firms that receive more information from their network are more likely to enter because a part of the informational cost burden related to entering the foreign market is offset by the network. We outline the limitations and strengths of this novel modeling approach and derive our estimation equation. An extended discussion of model identification is deferred to section 4.1.

[1] For an overview of the role of networks to trade see Chaney (2016).

[2] Notable exceptions are Connell et al. (2019) who relate export entry decisions to interactions with wholesalers and Patault and Lenoir (2021) who use labor movements between firms to explain subsequent export decisions of the hiring firm.

2.1 Stylized entry framework

We follow the framework of Koenig (2009) to study the decision of firm i to enter foreign market d. In this stylized setting, firms have absolute certainty about their expected profits $\Pi_{id,t}$ but face sunk entry costs f_d whenever entering a foreign market. A firm will start to export if the present value of profits (assuming constant discount factor r) exceeds the costs of entry. The probability to enter market d is thus

$$Pr(y_{id,t} = 1) = Pr\left(\frac{\Pi_{id,t}}{r} > f_d\right) \tag{1}$$

Firms therefore face a trade-off between costs and benefits of exporting to the foreign market and their ultimate decision will rest on the relative strength of both elements. Suppressing time subscript t, firm profits in market d can be described as

$$\Pi_{id} = p_{id}q_{id} - a_i w_i q_{id}$$

The first term on the right-hand side represents firm sales as the product of price p_{id} and demand q_{id} in the foreign market whereas the second describes the production costs which per unit of demanded quantity q_{id} amount to $w_i a_i$, the product of nominal wages w_i and inverse productivity a_i[3].

The model relies on a canonical setting where single product firms operate under monopolistic competition and consumers have CES utility which means the demand for firm i's products in market d is given by $q_{id} = p_{id}^{-\sigma} P_d^{\sigma-1} \mu_d E_d$ where $P_d^{\sigma-1} = \left[\int_l p_{ld}^{1-\sigma} dl\right]^{\frac{1}{1-\sigma}}$ represents the price index in market d, σ is the elasticity of substitution, μ_d is the expenditure share devoted to the representative industry and E_d denotes the level of income in d. The optimal mill price charged by firm i in this setting is $p_i = \frac{\sigma}{\sigma-1} a_i w_i$ as a constant markup over marginal costs $a_i w_i$. The final price faced by foreign consumers is $p_{id} = p_i \tau_d$ where τ_d represents ad-valorem iceberg-type trade costs related to shipping goods to market d. The profit of firm i in foreign market d is therefore

$$\Pi_{id} = \left[\frac{a_i w_i \tau_d}{(\sigma-1)P_d^{1-\sigma}}\right]^{1-\sigma} \mu_d E_d \tag{2}$$

Plugging Π_{id} into 1 and assuming (for now) that entry barriers f_d are common to all prospective entrants we can express a firm's entry decision as

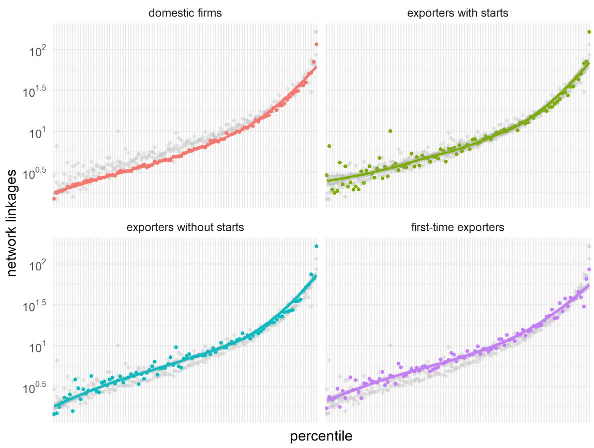
$$Pr(y_{id,t} = 1) = Pr\left(\left[\frac{a_{i,t} w_{i,t} \tau_{d,t}}{(\sigma-1)P_{d,t}^{1-\sigma}}\right]^{1-\sigma} \frac{\mu_{d,t} E_{d,t}}{r} - f_d > 0\right) \tag{3}$$

There are several things worth noting at this point. First, equation 3 illustrates that in the canonical setting all sources of firm heterogeneity are generated from firm-level characteristics, namely wages $w_{i,t}$ and (inverse) productivity $a_{i,t}$. This is emblematic of Melitz (2003) type trade models in which firm productivity is the key determinant in explaining the sorting of firms into exporters and non-exporters. Similarly, the Koenig (2009) model predicts that firms with higher productivity (lower $a_{i,t}$) are more likely to start exporting in presence of common entry costs f_d. Second, it highlights the role of gravity for the extensive margin of trade. Destinations with

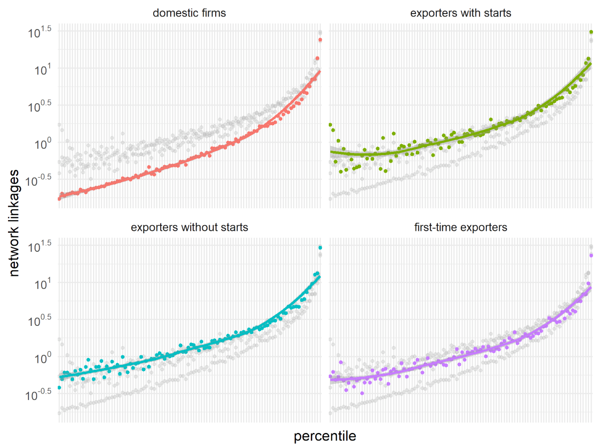
[3] The units of labor needed to produce one unit of q_i.

higher income levels $E_{d,t}$ or lower trade costs $\tau_{d,t}$ will attract more exporters conditional on firm characteristics.
Our goal is to extend this framework by a new source of firm heterogeneity that accounts for the fact that each individual firm has a unique network which may serve as a conduit for export-related information. Before formally introducing networks to the model in the next section, we provide some descriptive evidence to emphasize that network heterogeneity is not simply a primitive of firm productivity but contains additional variation that might be directly related to a firm's export participation.

Figure 1: Seller linkages by TFP percentile



(a) all linkages



(b) linkages to exporters

This figure shows shows the average number of buyers for a seller in a given productivity percentile. Seller productivity is computed using the approach of Levinsohn and Petrin (2003). Sellers are separated into four types: Non-exporters (red), exporters with export starts (green), exporters without starts (blue) and first-time exporters (purple). Figure 1a plots the average linkages to any buyer, while figure 1b plots the average linkages to buyers that export. The figure uses production network data of Belgian firms explained in detail in section 3.1.

Figure 1 relies on Belgian firm-level data explained in detail in section 3.1 and plots the average number of domestic buyer-seller linkages for sellers in a given productivity percentile. Both consist of four panels, each plotting the number of linkages for one seller type in color and the corresponding linkages of the other three seller types in grey. The only difference between the two graphs is the type of network linkages plotted on the y-axis. Figure 1a considers linkages to all buyers while figure 1b only considers the subset of linkages which involve exporting buyers.
Focusing on figure 1a, we observe two patterns. Firstly, across seller types, the total number of linkages seems to increase in seller productivity. This pattern is common to production networks (Zi and Bernard, 2021). High productivity sellers are likely to attract more buyers because they can charge lower prices than their competitors. Secondly, when comparing colored and grey dots in each panel, we see that conditional on seller productivity network sizes hardly differ across seller types. Comparing the network size of domestic (red) and exporting sellers (green, blue and purple) across productivity percentiles, we see that the network size premium of exporting sellers amounts to 28-37%. This indicates that exporting sellers overall seem to have more total network interactions than domestic sellers even after controlling for firm productivity.
This difference increases dramatically in figure 1b where we exclusively focus on seller linkages to exporting buyers. Network interactions with exporters are potentially important for a firm's

own export entry decision because they capture to what degree a firm is exposed to the export experience of other firms . What we see in the graph is that domestic sellers interact with fewer exporting buyers across productivity percentiles and that the network size premium of exporting sellers increases to 75-144%. This significant wedge in network interactions between domestic and exporting sellers cannot be explained by seller productivity (which we condition on) or total network size (as seen in figure 1a) and holds for incumbent and first-time exporting sellers alike. This strong correlation between seller's export status and their access to export experience through network linkages has important implications. First, it highlights that network interactions represent a new dimension of firm heterogeneity and are not simply a primitive of firm productivity. Exporting and non-exporting sellers with equal productivity interact with a different number of firms and this difference is especially pronounced when considering interactions with exporting buyers. Secondly, it introduces the possibility that networks are causing sellers to become exporters in the first place. To investigate this channel, we will formally introduce network interactions to the standard export entry framework outlined above and discuss how this additional dimension of firm heterogeneity can affect firm's entry behavior.

2.2 Augmented entry framework with network interactions

We treat each firm i's domestic production network as a pool of export-related information the firm can draw from. Each time a firm j in the network starts to export to a new destination d for the first time, it needs to pay sunk entry cost f_d which includes expenses to assess local demand preferences, search costs to identify foreign retailers and administrative costs related to obtaining mandatory certification for imported products or other non-tariff barriers. Each of these cost components involve a non-negligible share of information which the firm gathers upon entry and given sufficient interaction strength diffuses along network linkages. We define the diffusion triggered by each individual export start of firm j as an export signal y_{jd} which contains entry-related information for one particular market d. Information is assumed to diffuse immediately[4] upon entry of firm j. We formally model the impact of network interactions as an externality which enters firm i's entry equation 1, expressed in logs

$$Pr(y_{id,t} = 1) = Pr(\ln \Pi_{id,t} > \ln r + \ln f_d) \tag{4}$$

via entry cost term $\ln f_d$. This is achieved by making a functional form assumption which is inspired by the literature on social networks (Bramoullé et al., 2009; Calvó-Armengol et al., 2009) and explicitly expresses entry costs as a function of firm i's individual network

$$\ln f_{id,t} = \alpha_d - \delta \sum_j \bar{s}_{ij,t} x_{j,t} - \beta \sum_j s_{ij,t} y_{jd,t} - \psi_{S_{i,t}} + \varepsilon_{id,t} \tag{5}$$

where α_d denote sunk entry costs incurred in market d, $x_{j,t}$ and $y_{jd,t}$ represent time-varying characteristics. Export start decisions of network peers, $s_{ij,t}$ ($\bar{s}_{ij,t}$) are elements of a (row-normalized) binary interaction matrix S_t[5] which captures all domestic firm-to-firm interactions in the economy in year t, $\psi_{S_{i,t}}$ denote common shocks to all firms in firm i's network S and $\varepsilon_{id,t}$ is an

[4] Our empirical specification relies on annual firm-to-firm interactions. Immediate diffusion in this context means signals reach a connected firm i within the same calendar year as the export start of firm j.

[5] We will describe interaction matrix S_t in more detail in section 3.1.

idiosyncratic error term.
In this setting, network interactions can affect firm i's entry decision in two distinct ways. First, we allow entry costs f_{id} to directly respond to the average characteristics x_j of firms in the network[6]. This channel controls for network effects unrelated to the diffusion of export information such as productivity spillovers. Second, we allow entry costs f_{id} to directly respond to the total number of export signals y_{jd} received from the network. Each incoming export signal is assumed to lower entry barriers of firm i to destination d because the obtained information offsets a part of the informational cost otherwise incurred for export entry. An important distinction between both network externalities is the observational unit they operate at. Network characteristics δ vary at the firm level and can therefore only impact entry costs across export destinations (f_i). Network signals β however vary at the firm-destination level which means signals only affect entry costs of the market they originate from (f_{id}). Plugging equation 5 and 2 into equation 4 then yields our final entry equation

$$Pr(y_{id,t} = 1) = Pr\left(\boldsymbol{x_{id,t}}\gamma + \delta \sum_j \bar{s}_{ij,t} x_{j,t} + \beta \sum_j s_{ij,t} y_{jd,t} + \psi_{S_{i,t}} - \alpha_d - \varepsilon_{id,t} > 0\right) \tag{6}$$

where $\boldsymbol{x_{id,t}}$ captures firm and destination variables[7] related to firm profits in foreign market d. This equation closely resembles models from the peer effects (Manski, 1993; Bramoullé et al., 2009) and spatial economics (Anselin et al., 2008; Qu and Lee, 2015) literature where it is commonly referred to as a spatial autoregressive (SAR) model with panel data. Both strands guide our identification strategy which we defer to section 4.1. For the trade literature this explicit introduction of network interactions to a baseline model of export entry represents a novel modeling approach which warrants a further discussion regarding its strengths and weaknesses.

2.3 Concept limitations

The augmented entry framework described above introduces a new approach to allow network-transmitted information to affect export participation decisions of firms. Both the design of information diffusion and the choice to model network effects as an externality to market access costs carry important implications which will be discussed in turn.
Our diffusion mechanism relies on variation in export activity of network peers. An export entry decision of a firm in the network triggers the transmission of export signals which immediately propagate to connected firms. The type of diffusion mechanism assumed here is therefore based on the flow of new information rather than the flow of accumulated information available in the network. Our choice to focus on new rather than accumulated export information has several advantages. First, it offers a clear timing for the diffusion mechanism by linking both the creation and propagation of export information to the moment of export starts. In a model where information is accumulated over time the exact moment of diffusion is unclear. Secondly, limiting attention to new information ensures that the transmitted information remains relevant for firms

[6] We follow the standard convention of the social network literature and assume sunk entry costs respond to average rather than aggregate characteristics in firm i's network. This is achieved by row-normalizing entries in interaction matrix S_t such that $\sum_j \bar{s}_{ij} = 1$. As we are assuming all $s_{ij} \in \{0, 1\}$ in our empirical setting, $\sum_j \bar{s}_{ij} x_j$ simply represents an unweighted average of characteristics of all firms j in the network of firm i.

[7] $\boldsymbol{x_{id,t}} = (1-\sigma)(\ln a_{i,t} + \ln w_{i,t} + \ln \tau_{d,t} - (\sigma - 1)\ln P_{d,t}^{1-\sigma}) + \ln \mu_{d,t} + \ln E_{d,t}$.

at the receiving end. Entry information accumulated in previous periods would require additional assumptions about information depreciation rates to differentiate between the value of old and recent information. Our approach avoids this issue[8] by assuming that relevant information diffuses quickly along network linkages which is reasonable in our empirical setting[9].

A drawback of our diffusion approach is that we do not allow export signals to differ in value depending on which firm emitted a signal to the network. If firm A receives one export signal for the same destination from firms B and C, both incoming signals are treated equally. Signal heterogeneity could in principle be introduced by weighting matrix elements s_{ij} according to some characteristic of B and C, but in lack of a theory-consistent weighting scheme we refrain from this option. Another important limitation of our diffusion approach is that it does not grant insights about the relevance of individual information cost components. Entry-related information costs can be related to the assessment of demand preferences of foreign consumers, the search for local distribution partners or compliance with import requirements and product standards. Our model is not equipped to investigate the relative importance of each of these cost components which likely differ across products and destinations and instead focuses on their combined impact on export participation.

A second important aspect of our augmented framework concerns the way network interactions are allowed to influence firm entry decisions. Choosing to introduce network effects in form of an externality to market access costs f implies that network interactions are not strategic. Any benefit arising from interactions with other firms is assumed to be unexpected and not allowed to alter the profit optimization in equation 2. Firms do not account for network behavior when setting optimal prices but can benefit from network interactions in form of lower entry costs f. Our framework therefore abstracts from models with network games (König et al., 2019) where optimal firm behavior is interdependent due to the presence of strategic complementarities or models with strategic network formation (Badev, 2021; Hsieh et al., 2020) where firms anticipate network effects when choosing which agents to interact with. The network effect we study in this paper relates to the transmission of information related to export entry. This type of information is quite specific and unlikely to have a direct impact on firm prices or act as a core determinant of network formation which means our framework remains appropriate despite the lack of strategic interaction. One concern related to network formation however remains. While firms may not choose business partners strategically with the intent to acquire information about foreign markets, the formation of networks remains highly endogenous as firms do not choose their business partners at random. We discuss the identification challenges arising from endogenous network formation in detail in section 4.3 and present a solution to the likely selection bias inherent in firm-to-firm interactions.

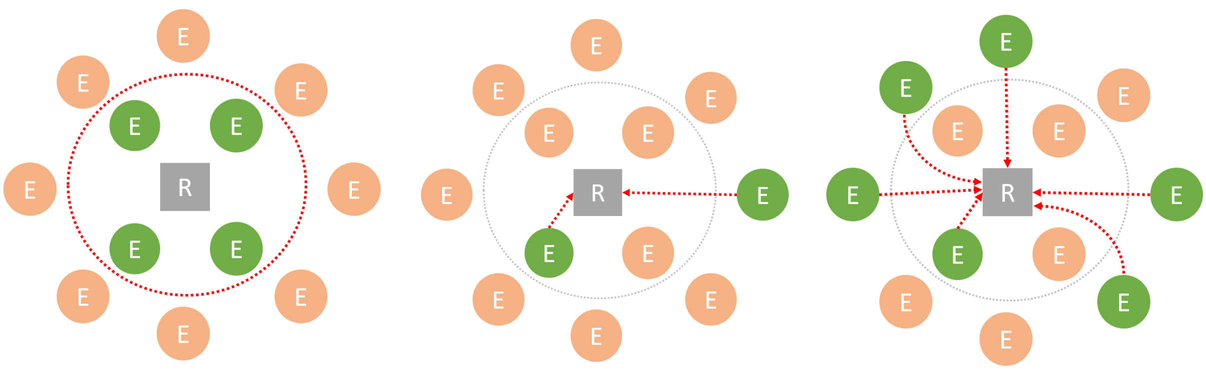
[8] Our approach does not allow for the value of information to depreciate before emission. Information depreciation however can still play a role after emission, if the receiving firm takes time to act upon the information received from its network.

[9] We capture network interactions and export starts at an annual level. This means our model assumes that export signals reach firms in the same year as they are emitted.

2.4 Concept advantages

Taken together, the diffusion mechanism and characterization of network effects as externalities augment the standard entry framework of Koenig (2009) by a new dimension of firm heterogeneity which offers several key advantages.

First, linking export participation to two distinct dimensions of firm heterogeneity offers new explanations for observed patterns at the extensive margin of trade. An important example is the existence of small exporters across export destinations documented by Eaton et al. (2011). A Melitz (2003) model with destination-specific entry costs f_d, constant marginal costs and productivity as the only dimension of firm heterogeneity typically struggles to explain the presence of small exporters if their productivity falls below model implied minimum thresholds required for export participation. Arkolakis (2010) addresses this shortcoming by presenting a model in which firms need to pay marketing costs to reach foreign buyers. Market access costs increase for each additional buyer the firm wants to reach, which explains the presence of small exporters since reaching a few customers is relatively inexpensive. Firm-specific network externalities studied in this paper offer an alternative explanation without assuming increasing market access costs. Firms at the margin that just fall below the minimum productivity threshold might still choose to export if they receive sufficient information from their network to bring the cost of entry below expected profits[10]. Allowing for both types of firm heterogeneity therefore captures additional variation which might be especially important to understand export participation decisions of small exporters.



(a) geographic proxy (b) labor movements (c) production network

Figure 2: Comparison of network approaches

Second, the augmented framework describes between-firm effects in a true network context which adds a substantial amount of empirical precision compared to the preceding export spillover literature. Prior studies like Koenig (2009) or Fernandes and Tang (2014) typically relate firms' export participation to the degree of export activity within the same geographic unit. This type

[10] Increasing market access cost in the Arkolakis (2010) model not only help to rationalize the existence of small exporters in presence of high entry costs but also match important export growth dynamics following episodes of trade liberalization. We do not investigate the relationship between network externalities and exporter dynamics post-entry in this paper. The comparison made therefore only seeks to illustrate how multi-dimensional firm heterogeneity could lead to new explanations for the presence of small exporters.

of geographic proxy is depicted in figure 2a and suffers from several important shortcomings as it implicitly assumes that all firms within the same geographic unit interact and that no interactions occur across units. These requirements directly contradict each other. To credibly capture firm-to-firm interactions in physical space, geographic units often need to be very small as spillovers may not even extend to firms across the street Bisztray et al. (2018). At the same time, a small radius increases the likelihood of missing interactions which occur across geographic units, making the optimal unit size an important but completely arbitrary determinant of diffusion. A related issue is the inability to capture different spillover effects for firms within the same unit as all of them by assumption share the same network. Both issues, the definition of unit size and lack of within-unit heterogeneity, similarly occur when capturing interactions via input-output tables[11] but are resolved in our augmented framework. The production network approach shown in figure 2c exploits all domestic firm-to-firm linkages revealing each firm's true network which is unique to each firm and free from any geographic constraints. These key characteristics allow us to move from the geographic group to the individual firm and express networks as a firm-specific conduit of information.

A related approach exploits labor movements to capture firm-to-firm interactions and is depicted in figure 2b. Choquette and Meinen (2015); Patault and Lenoir (2021) both investigate whether hiring employees from firms with export experience increases the likelihood to start exporting. While highly intuitive, this approach differs from ours in several important aspects. Hiring is more likely to be a strategic decision to access specialized information as firms pay particular attention to the skill set of applicants. Information therefore diffuses endogenously which stands in sharp contrast to the exogenous diffusion process in our setting. Moreover, the alternative approach describes a far more restrictive diffusion process as it rules out information transfers between firms that do not hire from each other. In contrast, we simply require firms to have business interactions which results in a much wider interaction network involving many more linkages and hence paths for information to flow between firms. To the extent that labor movements and business interactions overlap, we therefore generalize previous approaches by capturing a wider diffusion mechanism of export-related information[12].

A third important advantage of our augmented framework is the ability to contrast different behavioral assumptions related to how firms process incoming export signals. While networks expose firms to the export experience of their peers, not all network interactions will yield valuable information. If for example only a fraction of network peers export, most interactions

[11] Approaches using input-output data assume the same network for all firms operating in the same industry as interaction strength only varies at the sectoral level. This masks any network heterogeneity within sectors.

[12] Labor movements and business interactions will likely not overlap perfectly. While this means that our production network approach will likely miss certain linkages, it is captures a much wider extent of firm interactions than previous approaches. Patault and Lenoir (2021) for example use labor movements of sales managers as a proxy for firm-to-firm interactions. The implicit network size of their approach is therefore equivalent to the total number of competitors a firm hires from in any given year. As this number is not reported in their paper, we resort to an indirect comparison. In the extreme, the largest number of interactions from labor movements occurs if all French companies were to hire their entire staff from a distinct set of competitors every year. In this case, the average number of employees per firm serves as a proxy for the maximum network size as every employee is hired from a different competitor. Based on business demographics from Eurostat, the average number of employees per firm in France in 2019 is 5. When excluding firms with less than 10 employees, the average increases to 79. In contrast, The average number of buyers a seller interacts with each year in the Belgian production network is 20 for all sellers and 416 for sellers with at least 10 employees. Hence, even when taking labor movement-based network size to the extreme, production networks still capture a much larger number of firm-to-firm interactions.

will not transmit any export signals but may communicate other information unrelated to export entry. In the context of market access decisions in equation 1, non-export information represents noise as it does not increase firm profits or lower entry costs in the foreign market. The key question therefore is whether firms can filter export signals from noise and still make use of valuable information received from their network. Any stance taken on this question has direct implications for our functional form assumption on market access costs f. If we assume that unrelated information does not diminish a firm's ability to process incoming export signals, then network noise can effectively be ignored and we arrive at equation 5. We define this type of network externality, which relates entry costs to the total number export signals as *signal intensity*. In contrast, if we assume that firms' capacity to process network information is constrained, then non-export information matters as it will take up a part of firms' processing capacity. This reduces the impact of export signals as filtering out relevant information becomes costly. In other words, network externalities now depend on the share of relevant information out of total network information received. This changes our functional form assumption for entry costs f to

$$\ln f'_{id,t} = \alpha_d - \delta \sum_j \bar{s}_{ij,t} x_{j,t} - \beta \sum_j \bar{s}_{ij,t} y_{jd,t} - \psi_{S_{i,t}} + \varepsilon_{id,t} \tag{7}$$

Equation 7 is identical to equation 5 apart from a single change in the network information term β which now involves row-normalized interaction matrix elements $\bar{s}_{ij,t}$ to relate market access costs to the share of export signals $y_{jd,t}$ in total information received[13]. We define this type of network externality as *signal clarity* to underline the fact that network noise is explicitly accounted for. The distinction into signal intensity and signal clarity created by competing assumptions regarding the way firms process information is a key contribution of this paper and will become important to understand diverging network effects for firms with small and large networks which are explained in detail in section 3.2.3.

3 Data, empirical setting and descriptive statistics

To empirically investigate the role of network externalities for the extensive margin of trade we rely on detailed firm-level data from Belgium. As a small open economy Belgium represents an ideal setting for our analysis since it relies heavily on trade making export participation a key concern for firms and policy makers alike. In this section we first describe our main data sources and link them to the augmented framework derived above. We then present descriptive statistics to illustrate how network heterogeneity shapes the diffusion of export information among Belgian firms.

[13] To see this, recall that $s_{ij,t} \in \{0,1\}$ and row-normalized $\bar{s}_{ij,t} = \frac{s_{ij,t}}{\sum_j s_{ij,t}}$ such that $\sum_j \bar{s}_{ij,t} = 1$. The term $\sum_j \bar{s}_{ij,t} y_{jd,t}$ hence equals the average number of network peers which emit export signals at time t. If no firms in the network emit signals it equals 0, it all firms in the network emit signals it equals 1.

3.1 Data sources and sample selection

At the center of our analysis are three administrative datasets which are linked via unique firm identifiers and capture characteristics, export behavior and network interactions of Belgian firms for the years 2002 – 2014. Firstly, we use the Annual Account Filings database (National Bank of Belgium, 2002–2014a) which collects balance sheet information such as sales, revenues, input costs (labor, capital, material), 4-digit industry codes (NACE), zip code and ownership information from mandatory annual account filings of all firms operating in Belgium. We complement firm characteristics with annual import and export transaction data at HS6 product-level from the International Trade Dataset (National Bank of Belgium, 2002–2014b) which combines information from customs records and intra-EU trade declarations[14]. Together, balance sheet and trade data provide a detailed picture of performance and export activity of Belgian firms but do not grant any insights into firm-to-firm interactions. To fill this gap, we use the Business-to-Business Transactions Dataset (National Bank of Belgium, 2002–2014c) which records any buyer-seller transaction of firms operating in Belgium, provided the annual transaction value amounts to at least 250 €[15]. Belgian firms are required by law to file a breakdown of their annual sales by each individual buyer which allows us to identify individual firms involved in each transaction and thereby capture virtually all firm-to-firm interactions at an annual interval. To handle the vast amount of information contained in the combined dataset we implement important sample restrictions along firm, destination and network dimensions.

At the firm level, we follow the sample selection procedure of Dhyne et al. (2021) which significantly reduces the sample size while remaining very close to aggregate national statistics. In a first step this involves exploiting ownership information to single out observations which have unique identifiers but ultimately relate to the same firm. Identifiers in the data are constructed from value-added tax (VAT) numbers and some firms choose to use multiple VAT numbers for tax or accounting purposes. We aggregate these entries to the level of the firm which reduces the number of observations by around 4%. The second step of the selection procedure was originally introduced by De Loecker et al. (2014) and restricts our sample to firms with at least one full-time employee, more than 100 € of tangible assets, positive total assets in at least one reported year and positive labor costs and output. This step alone excludes more than 80% of the remaining observations as many firms in the original data are one-person companies[16]. The remaining sample is identical to the one used in Dhyne et al. (2021), includes between 90k-100k firms per year and remains very close to aggregate statistics in terms of value added, gross output, exports, and imports[17].

At the destination level, we only consider market access decisions for destinations outside the European Economic Area (EEA) as information frictions are expected to represent a much larger barrier to entry compared to highly integrated EEA countries[18]. Non-EEA destinations on aver-

[14] Intra- and extra-EU transactions have different reporting thresholds which are explained in appendix B.1

[15] For a detailed description of the dataset we refer to Dhyne et al. (2015)

[16] In 2012 there are 750,100 firms reporting less than 1 full-time employee.

[17] For a detailed comparison with aggregate statistics we refer to Table 1 in Dhyne et al. (2021)

[18] The list of EEA countries includes Greece, Lithuania, Portugal, Bulgaria, Spain, Luxembourg, Romania, Czech Republic, Hungary, Slovenia, Denmark, Croatia, Malta, Slovakia, Germany, Italy, the Netherlands, Finland, Estonia, Cyprus, Austria, Sweden, Ireland, Latvia, Poland, the United Kingdom, Norway, Liechtenstein, Switzerland, Iceland. We disregard all export transactions of Belgian firms to any of these countries for all sample years.

age account for roughly two-thirds of all export starts of Belgian firms which means our sample still captures the majority of activity at the extensive margin of trade. We follow Koenig (2009) and define an export start as a transaction to a destination which has not been served by the firm in the previous two years. Resuming exports to a foreign market after a single year of inactivity therefore are not treated as export starts[19]. This ensures that sufficient time has passed for market conditions to change such that information costs again become a relevant barrier to entry[20]. For our sample this implies that all observations of the first two years are dropped reducing the sample timeframe to 2004-2014. Further, we only consider firm-destination pairs with at least one export start across years to facilitate comparability across different estimation approaches[21].
At the network level, we start by characterizing the main network components. A network is defined as a collection of nodes and edges which in our case are represented by firms and their business transactions. Transactions (edges) therefore link firms (nodes) to each other and the transaction value (edge weight) gives an indication about the respective strength of each network interaction. In production networks edges are always directed because each firm involved in a transaction either acts as a buyer or a seller. In our setting, we need to distinguish between two distinct types of direction. First, the flow of goods and services from sellers to buyers along the supply chain which we define as a forward linkage. Second, the flow of money for goods and services sent from buyers to sellers which we define as a backward linkage. This distinction is important because network externalities in principle could go in either direction. In this paper we focus on information diffusion along backward linkages meaning sellers learn from their buyers. This direction has been identified as the most relevant one for information diffusion by the preceding export (Choquette and Meinen, 2015) and productivity (Javorcik and Spatareanu, 2011) spillover literature and is favored in our empirical setting. Sellers typically do not care which buyers they sell to which creates little incentive for them to communicate export-related information along forward linkages. Buyers on the other hand care about their suppliers as their own performance depends on the quality of sourced inputs. They might therefore be more likely to share export-related information with their suppliers, if this helps suppliers to understand the input requirements needed to sell the final product in foreign markets. Information in our empirical setting therefore diffuses from buyers to sellers.
While this clearly denotes which firms emit and receive export signals, in practice it is unlikely that all buyer-seller interactions meaningfully contribute to the diffusion of information. Sellers which only account for a small share of buyer sourcing may receive no information if the small transaction size requires less communication with buyers or indicates a low level of importance attached to the input in buyer production. We therefore need to distinguish between relevant and non-relevant buyer-seller interactions and exclude non-relevant linkages from the network as they are unlikely to play a role for the diffusion of information. To do so we compute the share

[19] Note that this allows for restarts within firm-destination pairs. In practice only 11% of entries are restarts.

[20] We assume that firms gather entry-related information upon entry. Firms that reenter after a single year of inactivity still possess very recent entry information and could benefit from their previous experience. By enforcing a 2-year period of inactivity we assume that entry requirements, consumer preference and non-tariff barriers in the destination have sufficiently changed such that information again presents a barrier to entry.

[21] Logistic regressions require variation in the outcome variable. To facilitate a comparison with results from linear probability model, we require at least one export start within each firm-destination pair which ensures sufficient variation for logistic regressions and allow to use the same sample for both estimation methods.

of total buyer sourcing accounted for by individual suppliers as

$$\nu_{ij,t} = \frac{\kappa_{ij,t}}{\sum_j \kappa_{ij,t}}$$

where $\kappa_{ij,t}$ represents the value of annual transactions between seller i and buyer j in year t taken from transaction value matrix $_t$. An interaction is defined as relevant for diffusion if suppliers account for at least 1% of buyer sourcing . Interactions that account for less than 1% of buyer sourcing[22] are treated as irrelevant for information diffusion and are excluded from the sample[23]. Applying this rule to all entries of transaction value matrix $_t$ leads to a binary interaction matrix S_t[24] with elements

$$s_{ij,t} = \begin{cases} 1, & \nu_{ij,t} \geq 1\% \\ 0, & \text{otherwise} \end{cases}$$

Each row of matrix S_t contains all linkages of seller i and the row sum indicates the number of buyers j a seller interacts with each year. As customary, self-links are not allowed which means all diagonal elements s_{ii} are set to zero.

3.2 Descriptive statistics

After implementing firm, destination and network restrictions our final sample contains characteristics of around 98,000 firms, 25,000 export starts to 188 non-EEA destinations and more than one million firm-to-firm interactions per year. The combined data allows us to trace the diffusion of export signals along network linkages and relate it to the entry behavior of Belgian firms. To understand how each data source contributes to this analysis we present descriptive evidence about firm behavior at the extensive margin of trade, the prevalence of signal diffusion and the role of network structure for the diffusion process.

3.2.1 Extensive margin of trade

Figure 3 shows the geographic dispersion of non-EEA export starts of Belgian firms between 2004-2014. Export decisions follow the rules of gravity and mainly occur in markets that are attractive due to their large size or limited distance to Belgium. One exception is the concentration of export entry in the Democratic Republic of the Congo. As a former colony the country retains strong ties to Belgium which potentially facilitates market access for Belgian exporters. Another important

[22] TO COMPLETE Our empirical results are robust to alternative thresholds as demonstrated in section 5.2.

[23] Our network sample is also subject to the firm-level restrictions described above which exclude 52% of network linkages from the sample. Of the remaining interactions, non-relevant linkages account for 85% in number but only account for 8% of total buyer sourcing. The network restriction therefore retains the majority of sourcing value $\nu_{ij,t}$ which is our key indicator of diffusion probability and greatly facilitates the analysis by reducing the sample size.

[24] We do not differentiate between transactions beyond the 1% threshold. While sellers with larger $\nu_{ij,t}$ might benefit from closer ties with buyers and hence receive more information, incorporating this idea would require a weighting scheme to map interaction intensity to information diffusion. In lack of a clear weighting scheme we resort to a binary measure but acknowledge that our equal treatment of above threshold linkages might result in a downward bias for network externalities.

pattern shown in figure 8 in appendix A.1 is distribution of export entry across geographic regions. While large countries like the US individually still account for the largest number of export starts, figure 8 shows that more than two-thirds of non-EEA new market entries occur in Africa and Asia. As these blocks comprise a large number of countries with different import regulations, consumer preferences and local supply networks, we expect entry-related information costs for these destinations to be high. This emphasizes the role of network externalities as many Belgian firms might want to reach the large consumers base in these emerging markets but lack the ability to overcome entry barriers.

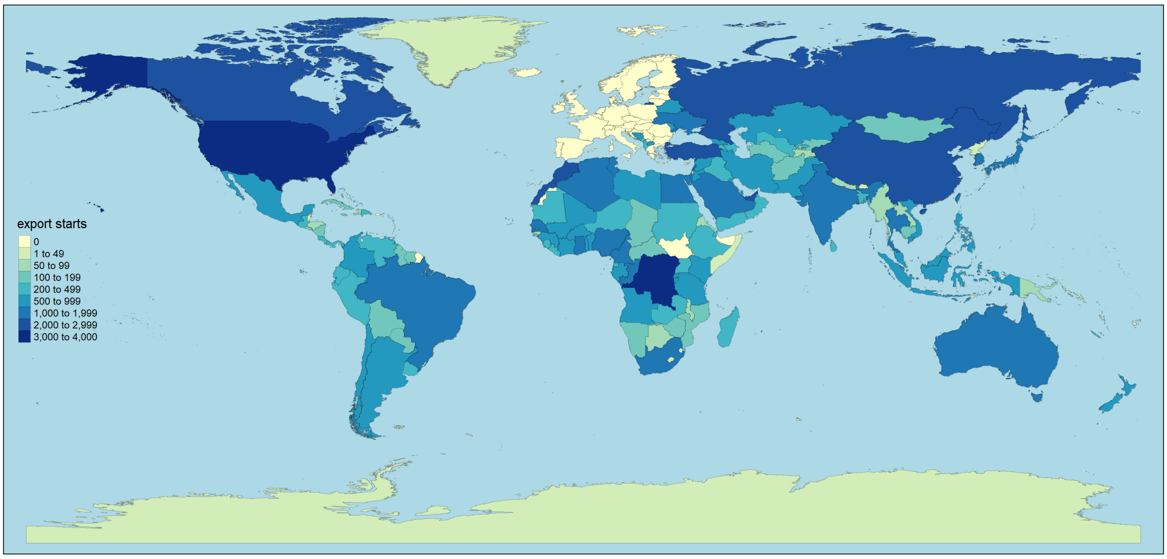


Figure 3: Geographic distribution of export starts (2004-2014)

3.2.2 Prevalence of export signals

Our augmented framework allows sellers to benefit from the export experience of their buyers in form of export signals that are emitted every time a connected buyer starts to export to a new destination. Our novel approach allows us to identify over 728,000 export signals received by sellers between 2004-2014 which highlights the vast scope of export information diffusing between firms. We differentiate export signals into matching and non-matching signals to indicate whether the origin of the incoming signal matches the destination the seller starts exporting to. Matching signals therefore represent valuable information information related to a firm's own export entry whereas non-matching signals capture the overall scope of the diffusion process.
Table 1 illustrates the prevalence of both signal types in each year and over the entire sample period. A first insight is that despite the large number of signals identified, a majority of sellers do not receive any export signals. Each year only 5.5% of sellers benefit from matching export signals which emphasizes that many entry decisions are still taken in absence of network externalities. This means there remains a large amount of cross-sectional variation we can exploit for our empirical analysis[25]. Second, the distribution of firms receiving export appears to be highly

[25]If most sellers received export signals in every period, identification of network externalities would mostly rely on within-firm variation in received export signals over time. Table **??** shows that our analysis can rely on both within- and between-firm variation when estimating network effects.

skewed. Over the entire sample period, roughly one half of all sellers do not receive any signals while a quarter of them receive more than 5[26]. This concentration of information diffusion among a small number of sellers is related to firm and network characteristics and further explored below.

Table 1: Share of firms receiving export signals

	per year		2004-2014	
signals	any signal	matching signal	any signal	matching signal
0	0.777	0.945	0.555	0.740
1	0.077	0.041	0.057	0.074
2	0.041	0.008	0.040	0.035
3	0.026	0.003	0.029	0.023
4	0.017	0.001	0.024	0.017
5	0.012	0.001	0.020	0.012
more than 5	0.045	0	0.248	0.081

This table indicates the share of firms that receive export signals in a single year and over the whole sample period. Matching signals represent the subset of total signals that originate in the same market as the subsequent export entry.

3.2.3 Network descriptives

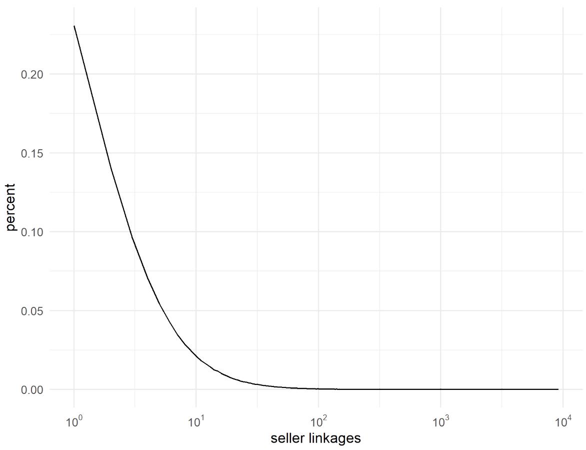


Figure 4: Distribution of seller linkages

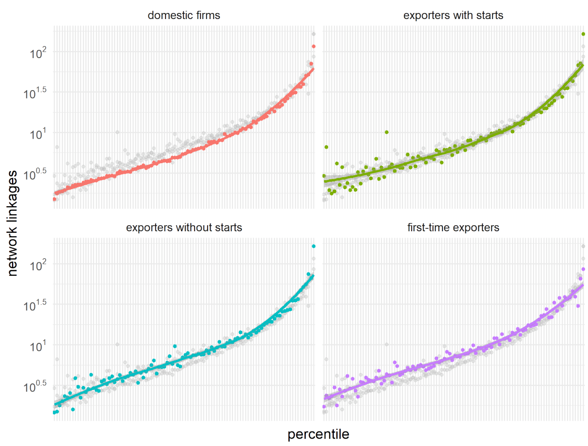


Figure 5: Seller linkages by sales percentile

We start with presenting a first set of characterisitcs related to network heterogeneity. Seller networks on average consist of 14 different buyers including 2 exporters and 1 export starter. These numbers mask a large amount of heterogeneity in firm networks as the distribution of seller linkages is highly skewed[27]. Figure 4 shows that while 25% of sellers only maintain a single network interaction, sellers in the top decile on average count over 1000 linkages to buyers each year. These vast differences in network size are closely related to seller size as shown in figure 5 . Sellers in higher sales percentiles (a proxy for firm size) on average interact with more buyers

[26] The full distribution of export signals is shown in appendix A.2.

[27] The distribution of linkages to exporters and export starters is equally skewed as shown in appendix A.3.

which holds for exporting and non-exporting sellers alike. This pattern is common to production networks (Zi and Bernard, 2021) and typically explained by superior firm performance. Large sellers can offer products at better quality or lower prices and thereby attract a larger number of buyers.

This complementarity between seller and network size has direct implications for the export participation pattern predicted by our augmented framework. If larger networks are associated with more export information and lower market access costs f_{id}, then network heterogeneity will simply reinforce selection patterns of the standard framework[28] by increasing the existing advantage of large sellers. In that case firm and network heterogeneity act as complements and our main contribution would be to estimate the relative importance of each channel. Conversely, if larger networks do not ensure more export information, then firm and network heterogeneity act as substitutes and each channel benefits firms of different size.

Figure 6: Seller network size and mean buyer export probability in 2014

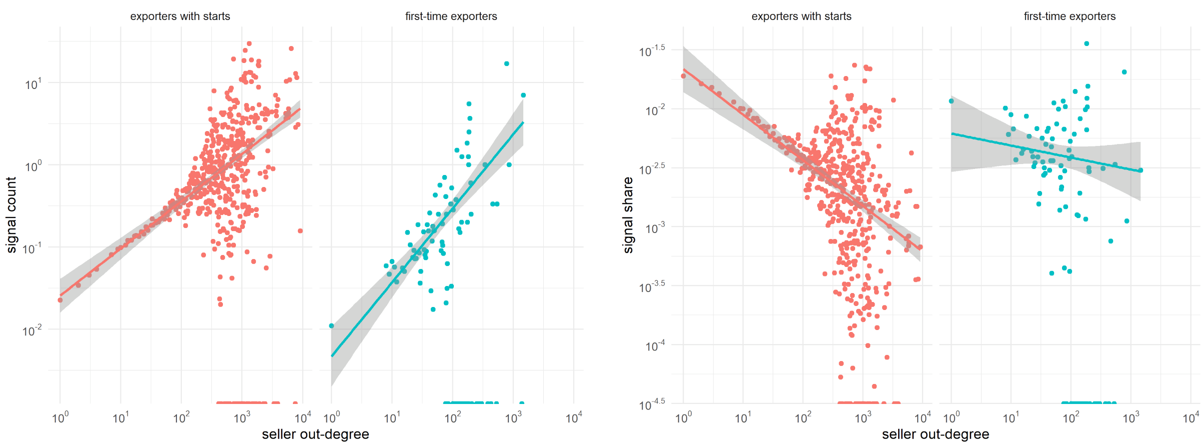
The latter case is of particular interest in our setting due to a second set of network characteristics related to negative degree assortativity. While larger sellers typically have larger networks, the average performance of their buyers is worse. This surprising fact can be seen as a curse of large sellers' own success. Being able to offer products at lower prices than competitors implies that even underperforming firms can afford these products which increases the total number of buyers but lowers average buyer performance. Small sellers in comparison tend to be less productive, offer products at relatively high prices which means only very performant firms are

[28]In absence of network interactions, larger (more productive) firms are already more likely to access foreign markets due to their superior firm performance (Melitz, 2003). This fact receives large empirical support as exporters tend to be larger and more productive than non-exporters (Bernard et al., 2003).

able to buy them. Figure 6 illustrates this pattern for seller networks in 2014 and shows how the share of exporting buyers changes with network size. The average buyer in a large network is significantly less likely to be an exporter compared to the average buyer in a small network. This relationship holds across years, seller types and alternative buyer characteristics such as sales, employment, productivity or export starts as shown in appendix A.4. If information diffusion depends on average rather than aggregate buyer behavior in the network, negative degree assortativity therefore reverses the impact of network heterogeneity on export entry. In this setting, small firms are expected to receive more export information than large firms as the average buyer is more likely to emit export signals.

3.2.4 Network externalities

A key takeaway from the previous section is that network externalities do not necessarily favor sellers with larger networks. Negative degree assortativity in production networks introduce a penalty that increases in network size and becomes important if the average network behavior is of interest. The decision to relate outcomes to aggregate or average network activity therefore has first-order importance for model predictions.
This finding is directly linked to the two competing behavioral assumptions of information uptake discussed in section 2.4. If we assume network noise can be ignored, then the aggregate number of export signals matter (signal intensity). If we instead assume network noise matters because firms' ability to process information is constrained, then we care about the relative number of export signals (signal clarity). Figure 7 shows how signal intensity and clarity change according to seller network size. As expected, we see a diverging pattern across measures. While signal intensity indicates that large firms receive more export-related information than small firms, the opposite is the case for signal clarity where network noise penalizes sellers with large networks. Choosing between signal intensity and signal clarity therefore not just entails different behavioral assumptions but also delivers divergent predictions for which group of sellers ultimately stands to gain from network externalities.



(a) signal intensity (b) signal clarity

Figure 7: Network externalities in 2014

4 Econometric framework

Introducing network interactions to the standard model of export entry creates several econometric challenges. Export decisions now not only depend on sellers' own characteristics but also on the characteristics and export experience of their buyers. We first outline the new identification challenges created in this augmented setting and then describe how we approach them in our empirical application.

4.1 Identification

As a starting point we rewrite our entry equation 6 in matrix notation. Using S_t and $\bar{S}_t$ to indicate standard and row-normalized binary interaction matrices we get

$$Pr(y_{d,t} = 1) = Pr(\boldsymbol{x_t}\gamma + \bar{S}_t\boldsymbol{x_t}\delta + \beta S_t y_{d,t} + \psi_{S_t} - \alpha_d + \varepsilon_t > 0) \tag{8a}$$

and

$$Pr(y_{d,t} = 1) = Pr(\boldsymbol{x_t}\gamma + \bar{S}_t\boldsymbol{x_t}\delta + \beta \bar{S}_t y_{d,t} + \psi_{S_t} - \alpha_d + \varepsilon_t > 0) \tag{8b}$$

for signal intensity and clarity respectively. Following the terminology of the peer effects literature these are commonly known as local-aggregate and local-average models. Both allow networks to affect seller outcomes in three distinct ways. First, in form of contextual peer effects which relate seller outcomes to buyer characteristics. These capture general externalities unrelated to export information such as productivity spillovers. Second, in form of endogenous peer effects which represent signal intensity or clarity. Third, in form of network fixed effect ψ_{S_t} which expose sellers to common shocks occurring in their immediate network. While this creates interesting interdependencies in firm behavior, the models cannot be readily taken to the data before dealing with several econometric challenges. To facilitate the discussion of each challenge, we take an intermediary step and assume $E(\varepsilon_t|\boldsymbol{x_t}, S_t) = 0$) which means networks form exogenously after conditioning on observable firm characteristics. We relax this assumption in section 4.3 where we introduce a network selection model to control for the potential selection bias created by endogenous network formation.

4.1.1 The reflection problem

A first challenge arises from the joint determination of buyer and seller outcomes y_t. If both terms enter the equations 8a and 8b contemporaneously, common shocks can lead to a simultaneity of buyer and seller export behavior. This well-known *reflection problem* (Manski, 1993) can prevent the separate identification of contextual and endogenous peer effects δ and β if individual firm networks do not sufficiently overlap[29]. In that case firm-to-firm linkages create separated network clusters in which firms only interact with members of the same cluster but have no linkages with firms in other clusters. If firms then experience a common shock, contextual and endogenous network effects become perfectly collinear as all firms within the same cluster act simultaneously

[29] In our setting, firm networks do not overlap if sellers act as exclusive suppliers for all buyers in their network and buyers themselves source but do not sell (= have positive indegree but zero outdegree).

and there is no variation from cross-cluster linkages to separately determine the impact each network channel.

The separate identification of both network effects is of particular importance in our setting, as we want to ensure that our main coefficient of interest β does not capture general spillover effects unrelated to export information. Bramoullé et al. (2009) and (Liu et al., 2014) show how this can be achieved in network settings for local-average and local-aggregate models respectively. In a local-average model contextual and endogenous networks are identified if identity matrix I, and interaction matrices S, and S^2 are linearly independent. In a local-aggregate model separate identification requires the rowsums of S to be non-constant and linear independence between I, S, $\bar{S}$ and $S\bar{S}$. Both sets of conditions are met in our setting as linkages in production networks are typically unidirectional which ensures linear independence of network matrices due to the presence of intransitive triads[30] and the fact that each seller interacts with a different number of of buyers leading to a non-constant rowsum of S.

Despite meeting the general conditions to identify network effects in a contemporaneous setting, we take a different approach because sellers are expected to respond to incoming export signals with delay. Assuming a temporal lag between signal reception and response is more realistic in our setting as sellers may take time to process information and adjust their production processes before entering a foreign market. Empirically, a delayed response also mitigates concerns related to the timing of buyer and seller export starts within the same year by ensuring that all sellers have sufficient time to react irrespective of the exact time a signal was received[31]. Lagging network effects changes our equations for signal intensity and signal clarity to

$$Pr(y_{d,t}=1) = Pr(\boldsymbol{x_t}\gamma + \bar{S}_{t-1}\boldsymbol{x_{t-1}}\delta + \beta S_{t-1}y_{d,t-1} + \psi_{S_{t-1}} - \alpha_d + \varepsilon_t > 0) \tag{9a}$$

and

$$Pr(y_{d,t}=1) = Pr(\boldsymbol{x_t}\gamma + \bar{S}_{t-1}\boldsymbol{x_{t-1}}\delta + \beta \bar{S}_{t-1}y_{d,t-1} + \psi_{S_{t-1}} - \alpha_d + \varepsilon_t > 0) \tag{9b}$$

Buyer and seller export starts now no longer occur simultaneously which resolves issues related to Manski's reflection problem. Conceptually, the change also marks a departure from the local-aggregate and local-average models of peer effects literature. Instead, it brings our approach closer to time-space recursive models[32] studied in spatial economics Anselin et al. (2008); Halleck Vega and Elhorst (2017) where current outcomes y_t are related to past network outcomes $\bar{S}_{t-1}y_{t-1}$. An important difference to time-space recursive models is that we do not consider lagged seller outcomes y_{t-1} as additional controls. This type of autocorrelation cannot occur in our setting due to the definition of export starts[33] which rules out entries to the same destination in two consecutive years.

A delayed response to network externalities facilitates model identification but requires additional assumptions regarding the timing of the underlying diffusion process. First, incoming

[30] An intransitive triad describes a network structure where firm A interacts with firm B, firm B interacts with firm C. It is called intransitive because there is no direct interaction between A and C.

[31] Sellers receive export signals at different points of the year. If responding to signals takes time, then sellers receiving a signal towards the end of the year are disadvantaged which may introduce a downward bias to the estimation of endogenous peer effects.

[32] Time-space here refers to two different types of lag from perspective of the dependent variable y_t. A spatial lag $S_t y_t$ indicating the relationship to network outcomes and a temporal lag $t-1$.

[33] An export start requires inactivity in the foreign market in the previous two periods. This implies that the two periods after an export start are excluded from the sample. A firm starting to export in year 3 and stopping in year 5 therefore is only faces entry decisions in periods 1,2,3 and 6. Years 4 and 5 are dropped from the sample.

export signals are not allowed to affect market access costs f_d in the same period. This assumption is compatible with a setting where networks form at the beginning of each period, firms choose whether to export and only then receive export signals. This order of events implies that information received in the current period only facilitates foreign market entry in next one. Second, to ensure conditional network exogeneity holds we need to assume that export signals received in period t do not influence network formation at the beginning of period $t+1$. This rules out that sellers form linkages strategically with the intent to reduce market access costs in the next period. Choosing to ignore export information received from previous linkages when forming new ones is restrictive but compatible with scenarios where the value of past signals is not yet realized at the beginning of the current period. Relaxing this assumption would require a formal model of strategic network formation as in (Badev, 2021) which is beyond the scope of this paper.

4.1.2 Correlated effects

A second challenge in our setting is to demonstrate that observed network effects are not driven by common network shocks. This becomes necessary because unobserved shocks affecting both buyer and seller outcomes create a *correlated effect* in buyer and seller behavior that is not related to network interactions. Correlated effects naturally arise in our study of export entry as firms face various shocks originating in both domestic and foreign markets. Failing to account for them will introduce a bias to estimated network coefficients δ and β and cast doubts on the relevance of network effects for export entry.

Turning to equations 9a and 9b, we see that assuming a lagged seller response to network effects rules out most correlated effects from temporary shocks as buyer and seller actions no longer occur in the same period. Lagged network effects however are no general remedy for correlated effects. Behavior of network members may still be correlated due to structural differences in domestic markets related to sectoral composition, local infrastructure and regional policy. We capture these channels with a combination of fixed effects and time-varying controls as explained in section 4.1.2. In practice, correlated effects from domestic markets however should only be a minor concern as they are unlikely to generate similar buyer and seller behavior for distinct export destinations. If all network members for example were located in the same city, they would equally benefit from positive shocks to local infrastructure which increases their likelihood to start exporting. While this would explain increased market access across export destinations and it fails to create buyer-seller correlation at the firm-destination level.

Common shocks originating in foreign markets on the other hand can cause correlated effects for distinct export destinations if firms respond to the same shock at different points in time. To illustrate this point, assume Chinese customs officials relax import requirements in period $t-1$ resulting in a decreased entry cost f_d. If buyers in the network immediately respond to the shock and start to export to China but sellers are only able to react to the shock in period t, the delayed shock response of sellers would be observationally indistinguishable from the network effect we try to capture. To isolate buyer-seller variation related to information diffusion, we need an instrument that can absorb any correlation in β created by the foreign shock.

We therefore instrument export starts of connected buyers $S_{i,t-1}y_{jd,t-1}$ with export starts occuring outside the seller's network in the same period. Denoting the set of firms unconnected to seller i in year t by matrix S^*, our instrument is formally defined as $S^*_{i,t-1}y_{jd,t-1}$. Export entry decisions of connected buyers and unconnected firms in the same period should be strongly correlated in the presence of common shocks originating in the foreign market. At the same time, they plausibly satisfy exclusion restrictions[34] as seller outcomes $y_{id,t}$ are not affected by firm behavior outside the network. To mitigate concerns related to misspecified seller networks, we exclude all firms from S^* which have an indirect connection to seller i through higher-order linkages or are connected through geographic proximity. This instrumentation strategy is currently being implemented and results will be shared in a future update of this draft (WORK IN PROGRESS).

4.2 Estimation

We now present our empirical framework. Under the assumptions discussed in the previous section and assuming networks form exogenously, we only need to make a distributional assumption for error terms $\varepsilon_{id,t}$ to take models 9a and 9b to the data. As a starting point, our benchmark estimation uses a linear probability model with fixed effects (LPM-FE) which assumes that errors $\varepsilon_{id,t}$ are $i.i.d$ and follow a normal distribution. A key advantage of the LPM-FE is the ability to easily accommodate high-dimensional fixed effects which allows us to capture unobservable time-invariant characteristics in domestic and foreign markets that otherwise may give rise to correlated effects. At the same time, assumed linearity limits the accuracy of predicted probabilities which can exceed the 0-1 interval.

Non-linear alternatives such as Probit and Logit models restrict predicted probabilities to the unit interval and therefore deliver more precise estimates for extreme values but typically suffer from an incidental parameter problem (IPP) when featuring high-dimensional fixed effects (Neyman and Scott, 1948). If the number of parameters that need to be estimated increases with sample size, maximum likelihood asymptotics no longer converge resulting in inconsistent parameter estimates. Our empirical setting is prone to this issue as the analysis considers export decisions at the firm-destination level which involves a large number of unobserved characteristics that need to be estimated. To evaluate benchmark estimates of LPM-FE model, we therefore contrast them with the fixed-effects logit estimator of Fernández-Val and Weidner (2016) and the fixed effects probit estimator of Hinz et al. (2021) which both feature a bias correction for the IPP while remaining directly comparable[35] to the LPM-FE via average partial effects.

Under normally distributed errors, we estimate the following reduced-form equation based on

[34] The critical assumption for this exclusion restriction to hold is that firms only influence each other via business interactions, meaning our sample selection accurately captures the relevant network of each seller. While this assumption is likely violated in practice due to unobserved firm linkages along other social dimensions, we expect our approach to perform reasonably well as most social networks are extremely sparse. Misspecified real firm linkages should therefore only represent a small fraction compared to the correctly identified absence of linkages between most firms.

[35] A common approach that avoids the IPP overall is the conditional logit model suggested by Chamberlain (1980). While delivering consistent parameter estimates it is not able to estimate average partial effects and therefore cannot be directly compared to the other methods.

our time-space recursive lag model:

$$Pr(y_{id,t} = 1) = Pr\left(\boldsymbol{x_{id,t}} + \sum_j \bar{s}_{ij,t-1}\boldsymbol{x_{j,t-1}}\delta + \beta\sum_j s_{ij,t-1}y_{jd,t-1} + \psi + \varepsilon_{id,t} > 0\right) \tag{10}$$

Seller export starts ($y_{id,t}$ are related to their own characteristics $\boldsymbol{x_{id,t}}$, network effects in form of buyer characteristics $\boldsymbol{x_{j,t-1}}$ and export signals $y_{jd,t-1}$ and a set of fixed effects ψ. We summarize the variables contained in each component below and present additional details in appendix B.2.

i. Seller characteristics $\boldsymbol{x_{id,t}}$ capture determinants that affect seller export decisions in absence of any network effects. These include firm-level controls such as total factor productivity (TFP), estimated using the procedure of Levinsohn and Petrin (2003), seller wages and employment-based seller size. Higher levels of TFP, wages and size are typically associated with increased export probability Bernard et al. (2003). Complementing these firm-level controls, we exploit available data about trade transactions to construct additional variables at the firm-destination level. First, we identify the products underlying a seller's export start and use this information to construct a firm-specific measure of import demand in each foreign market. This variable controls for export decisions as a direct response to foreign demand shocks. Second, we control for sellers' experience in a foreign market prior to their export start. Even without network linkages, sellers might accumulate expertise about destinations from other activities. We therefore add dummy variables to control for seller experience from importing, exporting to bordering destinations or destinations with historic ties[36]. Lastly, we control for a seller's overall export expertise via the share of export sales in total sales.

ii. Buyer characteristics $\boldsymbol{x_{j,t-1}}$ capture general network spillovers that affect seller entry across export destinations. Their main function is to ensure that export signal coefficient β is identified from destination-specific variation in the network. We include buyer sales and TFP to control for general spillovers unrelated to entry information.

iii. Lastly, ψ contains firm, destination, sector (4-digit NACE code) and year fixed-effects to control for unobservable time-invariant characteristics related to export entry. Firm fixed-effects capture unobserved firm expertise in specific markets. Destination fixed-effects account for differences in trade policy across countries. Sector fixed effects control for unequal integration of Belgian industries in global supply chains and year fixed effects for aggregate business cycle fluctuations. Together, these fixed effects rule out a lot of factors that potentially introduce correlation in buyer and seller behavior in absence of network externalities.

[36] The sequence of entry decisions is not random. Firms tend to enter markets that are similar to previous destinations (Morales et al., 2019) creating spatially correlated entry patterns (Albornoz et al., 2012)

Table 2: Regression sample (firm-years)

Statistic	N	Min	Pctl(25)	Median	Mean	Pctl(75)	Max
employees	89,120	1.00	4.50	12.60	76.45	36.90	59,691.68
wage (k)	89,120	0.70	39.51	48.59	52.76	60.30	574.71
TFP (log)	89,120	3.39	12.80	13.56	13.69	14.47	21.05
border dummy	89,120	0	0	0	0.15	0.2	1
history dummy	89,120	0	0	0.6	0.52	0.9	1
import dummy	89,120	0	0	0	0.09	0.1	1
export sales share	89,120	0.00	0.00	0.05	0.21	0.38	1.00
export demand (mn)	89,120	0.00	0.01	0.05	6.02	0.30	11,328.10
mean buyer sales (mn)	89,120	0.00	1.14	2.62	30.56	7.85	47,125.25
mean buyer TFP (log)	89,120	2.78	12.00	12.49	12.66	13.15	20.97

This table shows firm characteristics of our final regression sample. All variables have been aggregated to the firm-year level to facilitate readability. The reported number of observations therefore differs from the regression tables which capture entry decisions at the firm-destination-year level.

4.3 Network endogeneity

The preceding analysis relies on a conditional exogeneity assumption for network formation. As long as $E(\varepsilon_{id,t}|x_{id,t}, S_{i,t-1}) = 0$, interaction matrix elements $s_{ij,t-1}$ remain uncorrelated with outcome error $\varepsilon_{id,t}$ and network externality parameters δ and β can be accurately estimated. Network linkages however become endogenous if there are unobserved shocks which both affect network formation and export entry decisions. A shock that increases firms' search capacity for example could facilitate matching in both domestic and foreign markets. In that case, a positive relationship between export starts and network-mediated export signals would be caused by improved search capacity instead of information diffusion between firms. In consequence, interaction matrix S and outcome errors would be no longer uncorrelated and network effects estimated with bias.

To account for potential network endogeneity caused by unobserved shocks to linkage formation and export participation, we introduce the network selection model of Arduini et al. (2015) and Qu et al. (2017) to our estimation procedure. Formally, network endogeneity is modeled as a correlation between export entry error ε and network formation error ξ caused by an unobserved shock. To account for this correlation in our main entry equation, we estimate a dyadic network formation model and use predicted link probabilities to construct a selection correction term based on a Heckman-type mills ratio. Including this term as an additional variable in our outcome equation controls for the selection bias created by endogenous formation as adjusted outcome errors $\tilde{\varepsilon}$ are again uncorrelated with interaction matrix S. We are currently implementing this step and will share details of the procedure shortly. (WORK IN PROGRESS)

5 Preliminary results

In this section we use the empirical framework outlined above to assess the impact of network externalities for export participation decisions of Belgian firms. We proceed in several steps.

As a starting point, we obtain benchmark estimates of signal intensity and signal clarity from a LPM-FE, assuming no correlated effects from foreign shocks and exogenous networks. These estimates give a first indication about the relevance of network heterogeneity for export entry decisions and further demonstrate how behavioral assumptions of information uptake shape network effects. Next, we perform several robustness checks to ensure that benchmark results are not contingent on the specific model, timing assumptions, combination of fixed effects or sample selection. Finally, we will relax benchmark assumptions and allow for endogeneity in network effects from correlated effects and network formation. To account for correlated firm behavior due to unobserved shocks in foreign markets, we construct a network instrument and re-estimate our benchmark model using 2SLS. Endogenous network formation is addressed by introducing a dyadic network selection model which helps to eliminate the selection bias from correlated network formation and export behavior. Both endogeneity exercises are subject of ongoing research and not included in this preliminary draft. Results will be shared shortly. (WORK IN PROGRESS)

5.1 Benchmark results

Our first goal is to investigate whether network-mediated export signals indeed facilitate foreign market entry. An important advantage of out setting is the ability to identify all export signals that are received by a seller in a given year. This allows us to compare if signals from different origins have a different impact on sellers' subsequent export behavior. As entry requirements differ across markets, we expect export signals to only facilitate entry to the market the signal originates from. At the same time, a signal should not facilitate entry to a completely unrelated market. This key prediction of our augmented entry framework can be tested empirically by grouping incoming export signals into matching and non-matching signals based on whether signal origin and export destination overlap. Non-matching signals therefore capture export information that is received by sellers but should have no impact on subsequent export starts.

Table 3 summarizes the benchmark estimates of signal intensity while the full table is available in appendix C.1. Matching export signals (column 1) appear to be an important determinant for subsequent entry decision even after controlling for seller characteristics, experience in the foreign markets and general network spillovers related to peer characteristics. On average each additional signal received from the network increases the probability to start exporting by 0.47 percentage points. While we have seen in table 1 that the chance of receiving a matching signal is small, conditional on receiving one the economic significance of information diffusion can become sizeable if the number of received signals is sufficiently large. Moreover, the fact that matching signals turn out to be statistically significant even after controlling for seller productivity underlines that both firm and network heterogeneity matter to explain export participation patters of Belgian firms. In contrast, non-matching (column 2) and total export signals (column 3) seem to have no impact on entry behavior. This finding is important, as it shows that our diffusion mechanism is not picking up more general spillover effects that facilitate market access across destinations. Instead, entry barriers appear to be different in each market and as expected can only be reduced by matching signals.

In the remaining columns we disaggregate non-matching signals into three separate groups to corroborate these findings. Column 4 performs a placebo test by investigating the impact of

Table 3: Benchmark - signal intensity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0047^{***}					
	(0.0012)					
non-matching signals		-8.22×10^{-6}				
		(7.91×10^{-5})				
total signals			1.48×10^{-7}			
			(7.92×10^{-5})			
EEA signals				-3.5×10^{-5}		
				(0.0001)		
border signals					0.0002	
					(0.0004)	
history signals						0.0001^{**}
						(5.81×10^{-5})
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.123	0.123	0.123	0.123	0.123	0.123
Observations	473,494	473,494	473,494	473,494	473,494	473,494

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

received EEA signals. These destinations have been excluded from the sample in a previous step and signals should therefore under no circumstances have any impact on seller export starts which reassuringly is not the case. On the contrary, signals originating in markets with a close relation to the actual export destination could facilitate entry if entry barriers or local demand preferences are correlated in space. Columns 5 and 6 illustrate that there is mild evidence of these indirect channels. This result is interesting, as it shows that previous findings of spatially correlated entry patterns of firms' own export expansion (Albornoz et al., 2012; Morales et al., 2019) carries over to information diffusion. While direct information in form of matching signals is most valuable, indirect information from related markets to a lesser degree also facilitates entry.

Signal clarity results qualitatively mimic this pattern as shown in table 4 and appendix C.1. A higher share of matching signals (column 1) has the largest impact on entry behavior whereas non-matching signals only affect entry if they originate in related markets (columns 5 and 6) and remain insignificant if they they have no relation to the actual export destination whatsoever (column 4). To assess the economic relevance of signal clarity, recall that it measures the share of network interactions which yield valuable export signals. The coefficient for matching signal clarity of 0.0584 therefore indicates that if an additional 10% of network interactions were to emit a matching signal, this would on average increase the entry probability by 0.58 percentage points. Compared to signal intensity, this appears rather modest as the same impact is achieved from receiving a single matching signal. While the economic relevance of signal clarity may be

Table 4: Benchmark - signal clarity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0584***					
	(0.0129)					
non-matching signals		0.0074				
		(0.0051)				
total signals			0.0094*			
			(0.0051)			
EEA signals				0.0021		
				(0.0059)		
border signals					0.0155**	
					(0.0076)	
history signals						0.0182***
						(0.0046)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.123	0.123	0.123	0.123	0.123	0.123
Observations	473,494	473,494	473,494	473,494	473,494	473,494

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

smaller, its statistical significance carries an important implication. Noisy networks with low signal clarity might actually revert the positive effect of information diffusion as firms capacity to process information appears to be constrained.

To investigate this point further we repeat benchmark estimations of signal intensity and clarity for subsamples of large and small firms. The results are reported in tables 5 and 6. Even a simple separation of firms along median employment and sales reveal remarkable differences in network effects. Starting with signal intensity, the impact of matching signals is on average 3-4 times larger for small compared to large firms. Each additional export signal increases the export probability of small firms by 1.1-1.5 percentage points. Our benchmark estimates therefore mask significant differences in network effects by firm (and network) size. While large firms still benefit from incoming export signals, their weaker export response suggests that large networks limit the impact of externalities due to a redundancy of information or increased network noise. Results from signal clarity point towards the latter explanation as a marginal increase in signal clarity seems to disproportionately benefit large firms. As discussed in our descriptive analysis in section 3.2.4, network externalities do not seem to linearly increase in network size as information redundancy or network noise create diseconomies of scale. Contrary to firm heterogeneity in productivity which favors large and productive firms, network heterogeneity therefore might be particularly important for small firms as negative assortative matching tilts network externalities in their favor.

5.2 Benchmark robustness

To assess the validity of our benchmark results we perform several robustness checks. To begin with, we need to ensure that our results are not contingent on the chosen estimation approach. Benchmark estimates are derived from a LPM-FE whose linearity assumption implies that the marginal impact of matching signals is common to all firms and constant for each additional signal received. Non-linear alternatives such as Logit and Probit models relax this assumption and allow the marginal effects of export signals to vary with characteristics of each seller. At the same time, the inclusion of high dimensional fixed effects presents a real challenge for non-linear models as the asymptotics of the underlying ML estimator break down due to the incidental parameter problem (IPP). As the inclusion of fixed effects is essential to rule out unobserved heterogeneity that would otherwise plague our network estimates, we employ the bias-adjusted non-linear logit and probit models developed by Fernández-Val and Weidner (2016) and Hinz et al. (2021) which mitigate concerns related to the IPP and remain comparable to our linear estimates via APEs. Appendix C.2.1 shows the results from the direct comparison between linear and non-linear models for signal intensity and clarity. While coefficients naturally differ across models, APEs remain remarkably close to each other irrespective of which model is preferred. The comparison also reiterates the importance of network heterogeneity for export entry as matching signals remain statistically significant throughout all specifications.
A second set of robustness checks investigates to what extent our benchmark results depend on the imposed sample restrictions. The most important restriction in this regard is the definition of *relevant* network linkages. We only consider buyer-seller interactions if the seller accounts for at least 1% of total buyer sourcing. Empirically, this creates two distinct challenges. First, imposing a fixed interaction threshold might introduce measurement issues for seller networks. If a seller's sourcing share for a certain buyer fluctuates around the defined interaction threshold, we might miss valuable export information diffusing through this linkage as marginal decreases in seller's sourcing share would render the link irrelevant for our analysis. As a result, network effects might suffer from measurement bias since our linkage definition fails to capture export

Table 5: Firm size - signal intensity

Model:	empl $\leq$ median	empl $>$ median	sales $\leq$ median	sales $>$ median
Variables				
matching signals	0.0114^{***}	0.0040^{***}	0.0158^{***}	0.0040^{***}
	(0.0036)	(0.0011)	(0.0043)	(0.0011)
Peer characteristics	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes
firm FE	yes	yes	yes	yes
destination FE	yes	yes	yes	yes
year FE	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes
R^2	0.163	0.106	0.172	0.104
Observations	172,570	285,185	158,764	293,920

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 6: Firm size - signal clarity

Model:	empl ≤ median	empl > median	sales ≤ median	sales > median
Variables				
matching signals	0.0278*	0.1003***	0.0460**	0.0809***
	(0.0168)	(0.0205)	(0.0180)	(0.0197)
Peer characteristics	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes
firm FE	yes	yes	yes	yes
destination FE	yes	yes	yes	yes
year FE	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes
R^2	0.163	0.106	0.172	0.104
Observations	172,570	285,185	158,764	293,920

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

signals diffusing through linkages close to the interaction threshold. To assess this issue we split linkages into new and persistent ones based on whether an interaction was recorded as relevant in the previous year. We then recompute signal proxies for each linkage type. Measurement issues should only concern export signals from new linkages as persistent interactions by definition remain unaffected by fluctuations around the interaction threshold. Disaggregated results for new and persistent linkages are reported in appendix C.2.2. Reassuringly, both signals from new and persistent linkages positively affect export entry which emphasizes that our benchmark estimates are not solely driven by measurement error. Interestingly, signals transmitted through persistent linkages show a much stronger impact on entry than those transmitted through new linkages. While our diffusion approach does not differentiate between signal or interaction strength beyond the imposed interaction threshold, this result suggests that trust created from repeated interactions positively influences the diffusion of information in the network.

Another potential problem relates to the level of the interaction threshold itself. Choosing 1% as a universal cutoff value for relevant linkages is somewhat arbitrary which is why we need to ensure that benchmark results remain qualitatively unchanged for other cutoff values. We therefore repeat our benchmark estimations for a cutoff value of 5%. Intuitively, the higher value implies a stronger restriction for relevant interactions which results in fewer firm-to-firm linkages and a lower number of signals diffusing through the network. Results for the 5% interaction cutoff are presented in appendix C.2.3. Compared to our benchmark estimates, all key insights are mirrored at this alternative network definition. Entry decisions predominantly benefit from matching export signals while non-matching signals only mildly contribute to foreign market access if the originate from similar origins as the actual export destination. The same is true when contrasting estimates by firm size which reiterate that the marginal impact of an additional signal is stronger for small firms while large firms benefit from reduced network noise and higher signal clarity. While this shows that our main findings qualitatively do not depend on the chosen interaction threshold, the strength of network effects significantly differs across network samples. A stricter cutoff (5%) reduces the number of available export signals which become more valuable increasing the marginal effect of signal intensity. At the same time, it also reduces the overall number of linkages which mitigates the negative impact of network noise and thereby decreases the marginal effect of signal clarity. Different interaction cutoffs therefore equally underline the importance of network heterogeneity for the extensive margin of trade but when assessing marginal effect size researchers need to be mindful of the underlying cutoff value. Lastly, we contrast our benchmark results with a preceding literature of indirect exporting through wholesalers. Connell et al. (2019) show that interactions with wholesalers facilitate the internationalization of non-exporters in subsequent periods. Learning through wholesalers is treated as a channel to reduce uncertainty about consumer demand in the foreign market. After indirectly serving a destination through a wholesaler, firms become confident enough to enter the market directly. This mechanism is explicitly nested in our network approach as sellers can receive export signals from buyers that categorize as wholesalers. Importantly though, our approach is much more general as it allows firms to receive export signals from any interaction, including non-wholesalers. This motivates another disaggregation[37] of network effects into wholesaler and non-wholesaler export signals to assess if our general diffusion simply mirrors previous evidence. Results are reported in C.2.4 and indicate that while learning through wholesalers also facilitates a further internationalization of firms, non-wholesaler interactions appear to be much more important both in economic and statistical significance. Our results therefore corroborate previous evidence and complement it in an important way by showing that the information diffusion expand far beyond interactions with wholesalers.

6 Discussion

The preliminary results presented in the previous section indicate that networks act as an important determinant for export participation. Overcoming entry barriers therefore does not solely depend on firms' own characteristics but also on information transmitted through linkages in the domestic production network. This central finding has several important implications.
For trade theory, it means that both firm-level and network dimensions of firm heterogeneity need to be considered when describing behavior at the extensive margin of trade. Considering a setting with multidimensional firm heterogeneity relaxes the singular focus on firm productivity and helps to rationalize the existence of small exporters which under suitable network configurations face lower entry barriers and can therefore export despite insufficient levels of productivity. Future studies that seek to formally integrate network interactions into more elaborate trade frameworks should however be mindful about an important difference between both dimensions of firm heterogeneity. We find that network effects decrease in network size and relate this to increased levels of noise in large networks. This implies that positive effect of network heterogeneity is finite which is commonly not the case to firm-level heterogeneity. Very productive firms may still benefit from becoming more efficient whereas firms with large networks do not benefit from adding further linkages as network noise increases faster than the amount of valuable export signals. A formal model with multidimensional firm heterogeneity should therefore account for diseconomies of network scale.
On the policy side, our findings emphasize the role of information frictions in trade. Only match-

[37] We follow Connell et al. (2019) and identify wholesalers based on their 2-digit NACE code. Buyers with codes 46 and 47 are treated as wholesalers.

ing signals stimulate foreign market access which means that informational cost barriers differ substantially across export destinations. Any policy supporting firms to overcome entry barriers therefore needs to be adjusted to each individual export market. Our results suggest that networks provide a potentially powerful remedy for that problem. Policies that stimulate new interactions in domestic production networks would allows firms to draw on the experience of others and grant them access to a wide array of specialized export information. This could be an especially promising strategy to connect small and medium-sized enterprises (SMEs) to global markets as our results indicate that small firms stand to gain most from receiving additional export signals.

7 Conclusion

Export participation remains low across countries which causes concerns as it weakens competition in domestic markets (De Loecker and Warzynski, 2012) and restricts aggregate export growth (Eaton et al., 2009b). In this paper, we empirically study the determinants of export participation to provide insights on how firms can overcome trade barriers and serve foreign markets. A large preceding literature of heterogeneous firm trade models has emphasized the role of firm-level characteristics like productivity to rationalize observed entry patterns among firms.

We move beyond firm-level characteristics and investigate whether firm networks actively shape export entry decisions. Firms that interact with experienced exporters in their domestic production network receive export-related information which lowers sunk entry costs and thereby facilitates foreign market access. To formalize this mechanism, we introduce network interactions into a stylized model of export entry. The augmented model includes both firm-level and network heterogeneity and therefore allows to simultaneously assess the relevance of each determinant of export participation.

To estimate the model we rely on detailed data from the universe of Belgian firms which contains characteristics, cross-border trade transactions and business transactions in the domestic production network for every firm between 2002-2014. Combined, these unique datasets allow us to observe each firm's individual network as well as the export behavior of network peers. Every time a firm starts to export to a new export destination, it emits an export signal to connected firms which contains valuable entry information and lowers the market access costs. The number of received export signals varies across firms as each network is unique. This creates a second dimension of firm heterogeneity beyond firm productivity.

Taking this model to the data reveals that networks indeed play an important role in explaining export entry decisions of Belgian firms. According to our preliminary results, each additional export signal received from the network increases the entry probability to a specific foreign market by 0.5-1.4 percentage points. At the same time, having a large network is not always beneficial because a disproportionate number of interactions do not yield valuable export information. We relate this finding to negative assortative matching in the underlying network formation process. Large firms offer products at low prices and attract a large number of buyers, many of which are less productive themselves and therefore unlikely to emit export signals. Small firms in contrast offer products at higher prices which only productive buyers can afford. The average buyer in a small network is therefore more likely to emit export signals than the counterpart in a large network. In consequence, network effects are subject to diseconomies of scale.

Taken together, our findings demonstrate that network heterogeneity acts as an important new determinant of export participation beyond firm productivity. At the same time, they raise a number of important questions: Should we consider network effects as an externality or do firms instead select business partners strategically? How are the gains from trade distributed once network effects are taken into consideration? If network effects decrease in network size, is there an optimal number of linkages firms should maintain to maximize network benefits?

Answering these questions requires to link network formation to entry decisions in a more structural way and consider firm behavior in a general equilibrium model. Our results provide important empirical evidence for this future avenue of research and should promote a stronger consideration of networks in international trade.

References

Advani, A. and Malde, B. (2018). Methods to identify linear network models: a review. *Swiss Journal of Economics and Statistics*, 154(1):1–16.

Albornoz, F., Calvo Pardo, H. F., Corcos, G., and Ornelas, E. (2012). Sequential exporting. *Journal of International Economics*, 88(1):17–31.

Anselin, L., Gallo, J. L., and Jayet, H. (2008). Spatial Panel Econometrics. In Mátyás, L. and Sevestre, P., editors, *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Advanced Studies in Theoretical and Applied Econometrics, pages 625–660. Springer, Berlin, Heidelberg.

Arduini, T., Patacchini, E., and Rainone, E. (2015). Parametric and Semiparametric IV Estimation of Network Models with Selectivity. Technical Report 1509, Einaudi Institute for Economics and Finance (EIEF). EIEF Working Papers Series.

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

Arkolakis, C., Ganapati, S., and Muendler, M.-A. (2021). The Extensive Margin of Exporting Products: A Firm-Level Analysis. *American Economic Journal: Macroeconomics*, 13(4):182–245.

Badev, A. (2021). Nash Equilibria on (Un)Stable Networks. *Econometrica*, 89(3):1179–1206.

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

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

Bernard, A. B., Moxnes, A., and Ulltveit-Moe, K. H. (2018). Two-Sided Heterogeneity and Trade. *The Review of Economics and Statistics*, 100(3):424–439.

Bisztray, M., Koren, M., and Szeidl, A. (2018). Learning to import from your peers. *Journal of International Economics*, 115:242–258.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2020). Peer Effects in Networks: A Survey. *Annual Review of Economics*, 12(1):603–629.

Calvó-Armengol, A., Patacchini, E., and Zenou, Y. (2009). Peer Effects and Social Networks in Education. *Review of Economic Studies*, 76(4):1239–1267.

Chamberlain, G. (1980). Analysis of Covariance with Qualitative Data. *The Review of Economic Studies*, 47(1):225–238.

Chaney, T. (2016). Networks in International Trade. In Bramoullé, Y., Galeotti, A., and Rogers, B., editors, *The Oxford Handbook of the Economics of Networks*, pages 753–775. Oxford University Press.

Choquette, E. and Meinen, P. (2015). Export Spillovers: Opening the Black Box. *The World Economy*, 38(12):1912–1946.

Connell, W., Dhyne, E., and Vandenbussche, H. (2019). Learning about demand abroad from wholesalers: a B2B analysis. *National Bank of Belgium, Brussels*, NBB Working Paper(377).

De Loecker, J., Fuss, C., and Biesebroeck, J. V. (2014). International Competition and Firm Performance: Evidence from Belgium. *National Bank of Belgium, Brussels*, NBB Working Paper(269).

De Loecker, J. and Warzynski, F. (2012). Markups and Firm-Level Export Status. *American Economic Review*, 102(6):2437–2471.

Dhyne, E., Kikkawa, A. K., Mogstad, M., and Tintelnot, F. (2021). Trade and Domestic Production Networks. *The Review of Economic Studies*, 88(2):643–668.

Dhyne, E., Magerman, G., and Rubínova, S. (2015). The Belgian production network 2002-2012. *National Bank of Belgium, Brussels*, NBB Working Paper 288(288).

Eaton, J., Eslava, M., Kugler, M., and Tybout, J. R. (2009a). 8. Export Dynamics in Colombia: Firm-Level Evidence. In Helpman, E., Marin, D., and Verdier, T., editors, *The Organization of Firms in a Global Economy*, pages 231–272. Harvard University Press.

Eaton, J., Eslava, M., Kugler, M., and Tybout, J. R. (2009b). Export Dynamics in Colombia: Firm-Level Evidence. In *The Organization of Firms in a Global Economy*, pages 231–272. Harvard University Press.

Eaton, J., Jinkins, D., Tybout, J., and Xu, D. (2016). Two-sided Search in International Markets. *mimeo*.

Eaton, J., Kortum, S., and Kramarz, F. (2011). An Anatomy of International Trade: Evidence From French Firms. *Econometrica*, 79(5):1453–1498.

Fernandes, A. P. and Tang, H. (2014). Learning to export from neighbors. *Journal of International Economics*, 94(1):67–84.

Fernández-Val, I. and Weidner, M. (2016). Individual and time effects in nonlinear panel models with large N , T. *Journal of Econometrics*, 192(1):291–312.

Gaulier, G. and Zignago, S. (2010). Baci: International trade database at the product-level. the 1994-2007 version. Working Papers 2010-23, CEPII.

Halleck Vega, S. and Elhorst, J. P. (2017). Regional labour force participation across the European Union: a time–space recursive modelling approach with endogenous regressors. *Spatial Economic Analysis*, 12(2-3):138–160.

Hinz, J., Stammann, A., and Wanner, J. (2021). State Dependence and Unobserved Heterogeneity in the Extensive Margin of Trade. *arXiv:2004.12655 [econ]*.

Hsieh, C.-S., Lee, L.-F., and Boucher, V. (2020). Specification and estimation of network formation and network interaction models with the exponential probability distribution. *Quantitative Economics*, 11(4):1349–1390.

Javorcik, B. S. and Spatareanu, M. (2011). Does it matter where you come from? Vertical spillovers from foreign direct investment and the origin of investors. *Journal of Development Economics*, 96(1):126–138.

Koenig, P. (2009). Agglomeration and the export decisions of French firms. *Journal of Urban Economics*, 66(3):186–195.

König, M. D., Liu, X., and Zenou, Y. (2019). R&D Networks: Theory, Empirics, and Policy Implications. *The Review of Economics and Statistics*, 101(3):476–491.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *The Review of Economic Studies*, 70(2):317–341.

Liu, X., Patacchini, E., and Zenou, Y. (2014). Endogenous peer effects: local aggregate or local average? *Journal of Economic Behavior & Organization*, 103:39–59.

Manova, K. (2013). Credit Constraints, Heterogeneous Firms, and International Trade. *The Review of Economic Studies*, 80(2):711–744.

Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531–542.

Mayer, T. and Zignago, S. (2011). Notes on CEPII's Distances Measures: The GeoDist Database. *SSRN Electronic Journal*.

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

Morales, E., Sheu, G., and Zahler, A. (2019). Extended gravity. *The Review of economic studies*, 86(6):2668–2712.

National Bank of Belgium (2002-2014a). Annual Account Filings. Confidential Dataset.

National Bank of Belgium (2002-2014b). International Trade Dataset. Confidential Dataset.

National Bank of Belgium (2002-2014c). Business-to-Business Transactions Dataset. Confidential Dataset.

Neyman, J. and Scott, E. L. (1948). Consistent Estimates Based on Partially Consistent Observations. *Econometrica*, 16(1):1–32.

Patault, B. and Lenoir, C. (2021). How valuable are business networks? Evidence from sales managers in international markets. *mimeo*.

Qu, X. and Lee, L.-f. (2015). Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics*, 184(2):209–232.

Qu, X., Lee, L.-f., and Yu, J. (2017). QML estimation of spatial dynamic panel data models with endogenous time varying spatial weights matrices. *Journal of Econometrics*, 197(2):173–201.

Zi, Y. and Bernard, A. B. (2021). Sparse Production Networks. *mimeo.*

A Additional descriptives

A.1 Export starts

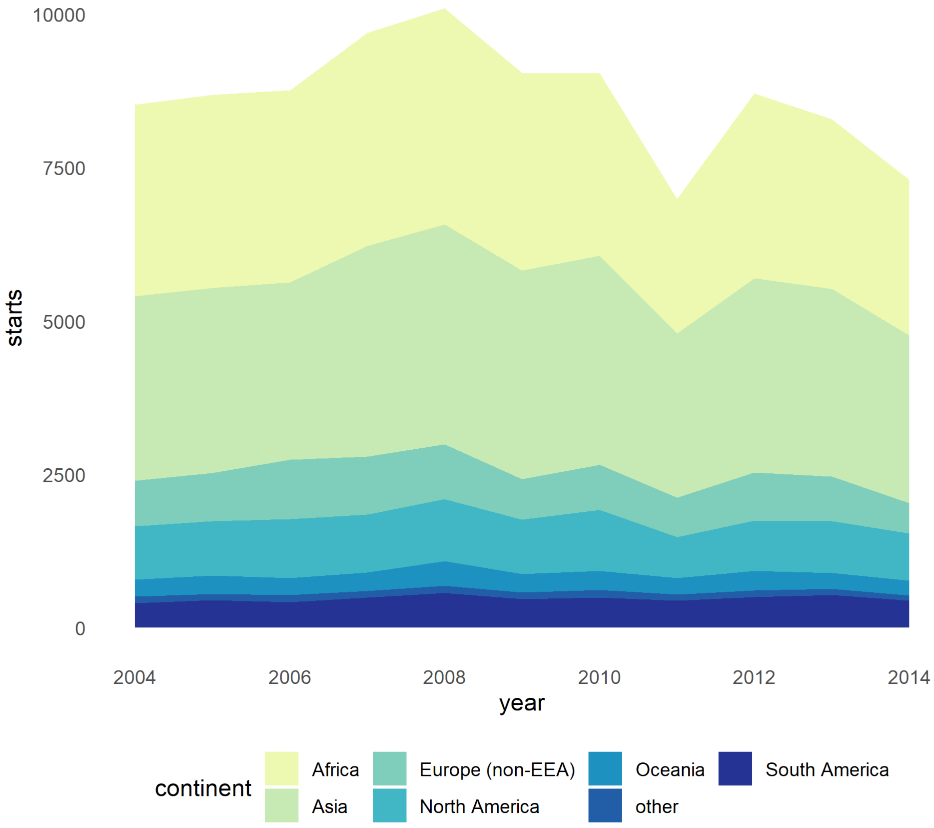
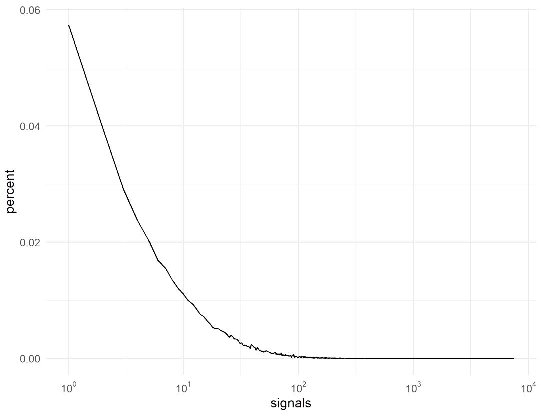
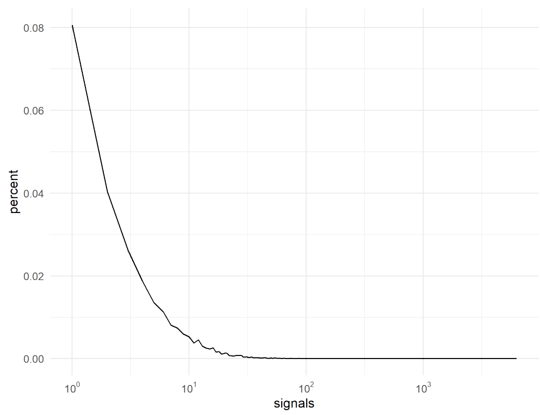


Figure 8: Share of non-EEA starts by region (2004-2014)

A.2 Export signals



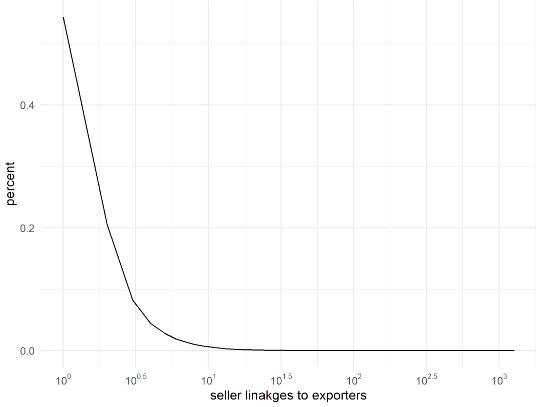
(a) any export signal



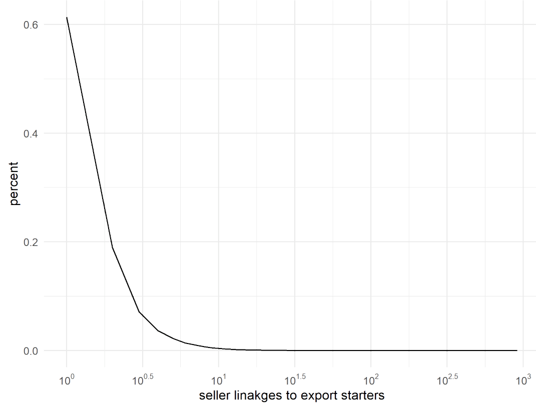
(b) matching export signal

Figure 9: Distribution of firms receiving export signals

A.3 Network linkages



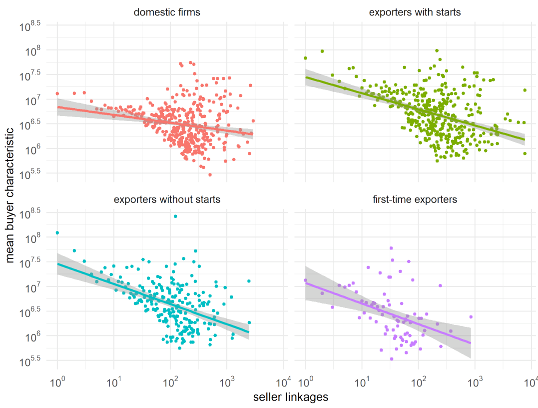
(a) linkages with exporters



(b) linkages with export starters

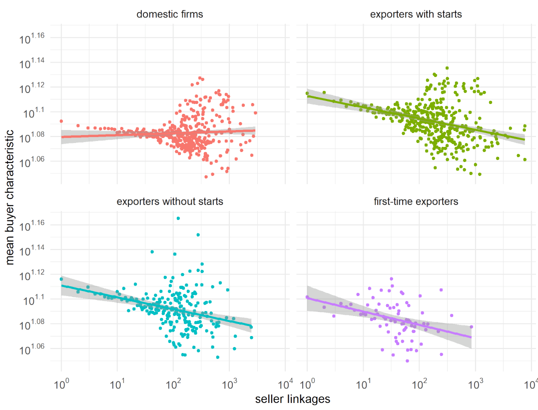
Figure 10: Distribution of seller linkages

A.4 Network assortativity

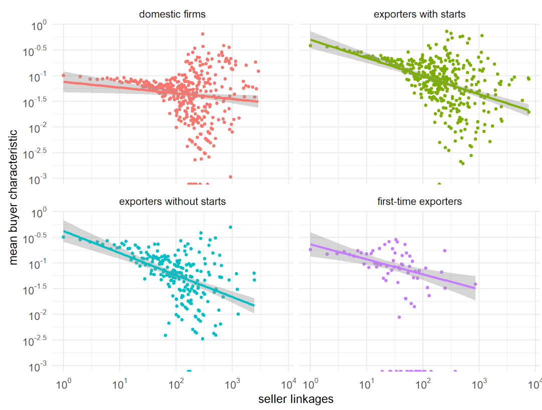


(a) average buyer sales

(b) average buyer employment



(c) average buyer productivity



(d) average buyer export start probability

Figure 11: Seller network size and mean buyer characteristics in 2014

B Additional dataset information

B.1 Reporting thresholds for trade transactions

The reporting thresholds differ across intra-EU and extra-EU trade transactions. Extra-EU export and import transactions follow a common reporting standard across all sample years. They are covered in the dataset if the transaction value exceeds 1,000 € or the volume is bigger than 1,000kg. In rare instances transactions below the minimum volume threshold are observed if the respective firm uses electronic reporting standards.
Intra-EU transaction thresholds are much higher and change over the sample period. Before 2006, they are reported if the combined import and export value of a firm exceeds 250,000 €. Between 2006-2010, the reporting threshold for imports was 400,000 € and 700,000 € for exports before both where harmonized to 700,000 € in 2010.
Our analysis mainly focuses on extra-EU transactions and therefore avoids measurement issues related to changing reporting thresholds or high threshold levels.

B.2 Construction of the regression sample

The variables used in our regression sample draw on the rich information contained in our merged dataset.

i. *Export starts* rely on detailed HS6 product-level export-transaction data which we aggregate to the firm-destination level. A firm with positive export transactions each year is counted as an exporter. An export start is defined as an export transaction to a destination that has not been served in the previous two periods. All observations within the two-year buffer period are dropped as firms by definition do not face an export decision. Likewise, non-starts are also only included in the data, if the firm has not been exporting in the past two years to ensure that a start could have occurred mean the firm faced a actual entry decision.

ii. Data on the number of *employees* and firm *wages* can be directly obtained from the available balance sheet data.

iii. *Total factor productivity* (TFP) is estimated using the approach of Levinsohn and Petrin (2003). The estimation requires data on firm sales, capital, labor and material inputs which are all available in the balance sheet data. Deflators for each input at 2-digit NACE codes are provided by the NBB based on internal price information. Our estimation is performed sector-by-sector and we only include sectors for which at least 50 non-missing observations are available.

iv. *Export experience dummies* rely on a combination of Belgian trade-transaction data for import and exports and the GeoDist database (Mayer and Zignago, 2011) freely available from CEPII's website. The latter includes information on bilateral relationships between all more than 200 countries including historic links and geographic borders. We merge this country relationship information with trade transaction data to create history and border dummies depending on the recorded relationship between Belgium and the respective trade partner. Import dummies on the other hand are only require the original trade transaction data and mark whether a seller has directly imported products from the future export destination. *Export sales shares* compare aggregate export values to sales information in the balance sheet records.

v. The idea for *export demand* is to capture the demand for the products underlying the export starts of Belgian firms in the foreign market prior to the actual export entry. To do so we proceed in several steps. First, we collect import data at HS6 product level for all destinations and sample years from the BACI database (Gaulier and Zignago, 2010) which we complement with WTO data for missing import information for Taiwan. Next, we identify the products underlying the export starts of each firm using the Belgian trade transaction database. For these products, we compute the export value at HS6 product-level in each destination originating in non-EEA countries. These non-EEA exports should capture changes in product demand in the destination without being correlated with Belgian exports due to common trade policy. For each firm, this gives us a proxy of how strongly their product was demanded in the destination prior to the export start. We then aggregate this export demand information to the firm-destination level and introduce it to the regression sample to control for the firm-specific export demand in each destination in each year t.

vi. Peer characteristics included in our regression sample are buyer *TFP* and buyer *sales* available from the Belgian balance sheet data. To relate buyer characteristics to sellers, we use row-normalized interaction matrices $\bar{S}_t$ and compute the average TFP and sales of buyers in a seller's network.

C Additional results

C.1 Benchmark regressions - full table

Table 7: Benchmark - signal intensity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0047***					
	(0.0012)					
non-matching signals		-8.22×10^{-6}				
		(7.91×10^{-5})				
total signals			1.48×10^{-7}			
			(7.92×10^{-5})			
EEA signals				-3.5×10^{-5}		
				(0.0001)		
border signals					0.0002	
					(0.0004)	
history signals						0.0001**
						(5.81×10^{-5})
log employment	0.0455***	0.0456***	0.0456***	0.0456***	0.0456***	0.0455***
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
log wage	0.0090	0.0089	0.0090	0.0089	0.0090	0.0090
	(0.0061)	(0.0061)	(0.0061)	(0.0061)	(0.0061)	(0.0061)
log TFP	0.0398***	0.0399***	0.0399***	0.0399***	0.0399***	0.0399***
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
log export demand	0.0179***	0.0180***	0.0180***	0.0180***	0.0180***	0.0180***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
border dummy	0.0686***	0.0687***	0.0687***	0.0687***	0.0686***	0.0687***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
history dummy	-0.0145***	-0.0145***	-0.0145***	-0.0145***	-0.0145***	-0.0145***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
export propensity	0.1698***	0.1698***	0.1698***	0.1698***	0.1698***	0.1698***
	(0.0113)	(0.0113)	(0.0113)	(0.0113)	(0.0113)	(0.0113)
import dummy	0.0986***	0.0988***	0.0988***	0.0988***	0.0988***	0.0988***
	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)
log peer size	-0.0007	-0.0005	-0.0005	-0.0005	-0.0006	-0.0006
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
peer TFP	7.14×10^{-5}	-5.84×10^{-5}	-4.87×10^{-5}	-6.74×10^{-5}	-3.48×10^{-5}	1.34×10^{-5}
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0028)
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.123	0.123	0.123	0.123	0.123	0.123
Observations	473,494	473,494	473,494	473,494	473,494	473,494

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 8: Benchmark - signal clarity

	export starts					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0584***					
	(0.0129)					
non-matching signals		0.0074				
		(0.0051)				
total signals			0.0094*			
			(0.0051)			
EEA signals				0.0021		
				(0.0059)		
border signals					0.0155**	
					(0.0076)	
history signals						0.0182***
						(0.0046)
log employment	0.0457***	0.0456***	0.0456***	0.0456***	0.0456***	0.0457***
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
log wage	0.0090	0.0089	0.0089	0.0090	0.0090	0.0089
	(0.0061)	(0.0061)	(0.0061)	(0.0061)	(0.0061)	(0.0061)
log TFP	0.0399***	0.0399***	0.0399***	0.0399***	0.0399***	0.0400***
	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0042)
log export demand	0.0179***	0.0180***	0.0180***	0.0180***	0.0180***	0.0180***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
border dummy	0.0686***	0.0686***	0.0686***	0.0687***	0.0685***	0.0686***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
history dummy	-0.0145***	-0.0145***	-0.0145***	-0.0145***	-0.0145***	-0.0145***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
export propensity	0.1699***	0.1698***	0.1698***	0.1698***	0.1699***	0.1698***
	(0.0113)	(0.0113)	(0.0113)	(0.0113)	(0.0113)	(0.0112)
import dummy	0.0988***	0.0988***	0.0988***	0.0988***	0.0988***	0.0988***
	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)
log peer size	-0.0006	-0.0006	-0.0006	-0.0005	-0.0006	-0.0006
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
peer TFP	-0.0003	-0.0008	-0.0011	-0.0002	-0.0002	-0.0010
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0028)
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.123	0.123	0.123	0.123	0.123	0.123
Observations	473,494	473,494	473,494	473,494	473,494	473,494

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.2 Benchmark robustness

C.2.1 Comparison of estimation methods

Table 9: Model comparison - signal intensity

	Dependent variable: export starts				
	(1)	(2)	(3)	(4)	(5)
Model	LPM-FE	Logit-FE	Logit-FE-IPP	Probit-FE	Probit-FE-IPP
Coefficient for *matching signal*	0.00488^{***}	0.0385^{***}	0.0384^{***}	0.0209^{***}	0.0208^{***}
	(0.00118)	(0.00595)	(0.00595)	(0.00338)	(0.00338)
APE for *matching signal*	0.00488^{***}	0.00558^{***}	0.00574^{***}	0.00531^{***}	0.00545^{***}
	(0.00118)	(0.000923)	(0.000924)	(0.000935)	(0.000936)
Fixed-effects					
firm FE	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes
Fit statistics					
Observations	487,516	477,263	477,263	477263	477263

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 10: Model comparison - signal intensity

	Dependent variable: export starts				
	(1)	(2)	(3)	(4)	(5)
Model	LPM-FE	Logit-FE	Logit-FE-IPP	Probit-FE	Probit-FE-IPP
Coefficient for *matching signal*	0.0596^{***}	0.354^{***}	0.342^{***}	0.214^{***}	0.208^{***}
	(0.0129)	(0.0727)	(0.0727)	(0.0424)	(0.0424)
APE for *matching signal*	0.0596^{***}	0.0512^{***}	0.0511^{***}	0.0544^{***}	0.0544^{***}
	(0.0129)	(0.0105)	(0.0105)	(0.0107)	(0.0107)
Fixed-effects					
firm FE	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes
Fit statistics					
Observations	474,676	464,551	464,551	464,551	464,551

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.2.2 Network persistence

Table 11: Network persistence - signal intensity

	export starts					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals (new links)	0.0032^{**}					
	(0.0014)					
matching signals (old links)	0.0073^{***}					
	(0.0023)					
non-matching signals (new links)		-0.0001				
		(0.0001)				
non-matching signals (old links)		0.0001				
		(9.99×10^{-5})				
total signals (new links)			-0.0001			
			(0.0001)			
total signals (old links)			0.0001			
			(9.98×10^{-5})			
EEA signals (new links)				-0.0003		
				(0.0002)		
EEA signals (old links)				0.0002		
				(0.0002)		
border signals (new links)					2.76×10^{-5}	
					(0.0006)	
border signals (old links)					0.0006	
					(0.0010)	
history signals (new links)						6.21×10^{-5}
						(0.0001)
history signals (old links)						0.0003
						(0.0002)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.123	0.123	0.123	0.123	0.123	0.123
Observations	473,494	473,494	473,494	473,494	473,494	473,494

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 12: Network persistence - signal clarity

	export starts					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals (new links)	0.0502***					
	(0.0164)					
matching signals (old links)	0.0704***					
	(0.0207)					
non-matching signals (new links)		0.0114**				
		(0.0057)				
non-matching signals (old links)		0.0024				
		(0.0067)				
total signals (new links)			0.0134**			
			(0.0057)			
total signals (old links)			0.0044			
			(0.0067)			
EEA signals (new links)				0.0008		
				(0.0065)		
EEA signals (old links)				0.0045		
				(0.0087)		
border signals (new links)					0.0169*	
					(0.0098)	
border signals (old links)					0.0131	
					(0.0118)	
history signals (new links)						0.0218***
						(0.0057)
history signals (old links)						0.0126*
						(0.0071)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
Fixed-effects						
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
Fit statistics						
R^2	0.123	0.123	0.123	0.123	0.123	0.123
Observations	473,494	473,494	473,494	473,494	473,494	473,494

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.2.3 Network threshold

Table 13: Network threshold 5% - signal intensity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0096***					
	(0.0022)					
non-matching signals		9.72×10^{-5}				
		(0.0003)				
total signals			0.0001			
			(0.0003)			
EEA signals				8.48×10^{-5}		
				(0.0005)		
border signals					0.0014	
					(0.0009)	
history signals						0.0004*
						(0.0002)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.119	0.119	0.119	0.119	0.119	0.119
Observations	362,517	362,517	362,517	362,517	362,517	362,517

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 14: Network threshold 5% - signal intensity - firm size

Model:	empl<=median	empl>median	sales<=median	sales>median
Variables				
matching signals	0.0174***	0.0083***	0.0199***	0.0084***
	(0.0066)	(0.0024)	(0.0074)	(0.0023)
Peer characteristics	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes
firm FE	yes	yes	yes	yes
destination FE	yes	yes	yes	yes
year FE	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes
R^2	0.162	0.100	0.167	0.100
Observations	132,664	218,406	123,372	224,269

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 15: Network threshold 5% - signal clarity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0269**					
	(0.0125)					
non-matching signals		0.0085*				
		(0.0047)				
total signals			0.0095**			
			(0.0047)			
EEA signals				0.0043		
				(0.0057)		
border signals					0.0040	
					(0.0077)	
history signals						0.0130***
						(0.0046)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.119	0.119	0.119	0.119	0.119	0.119
Observations	362,517	362,517	362,517	362,517	362,517	362,517

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 16: Network threshold 5% - signal clarity - firm size

Model:	empl<=median	empl>median	sales<=median	sales>median
Variables				
matching signals	0.0022	0.0537***	0.0148	0.0487***
	(0.0163)	(0.0199)	(0.0171)	(0.0185)
Peer characteristics	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes
firm FE	yes	yes	yes	yes
destination FE	yes	yes	yes	yes
year FE	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes
R^2	0.162	0.100	0.167	0.100
Observations	132,664	218,406	123,372	224,269

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.2.4 Wholesalers

Table 17: Wholesaler linkages - signal intensity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals (wholesaler)	0.0033*					
	(0.0020)					
matching signals (normal)	0.0054**					
	(0.0026)					
non-matching signals (wholesaler)		0.0002				
		(0.0005)				
non-matching signals (normal)		-0.0004				
		(0.0008)				
total signals (wholesaler)			0.0002			
			(0.0005)			
total signals (normal)			-0.0004			
			(0.0008)			
EEA signals (wholesaler)				0.0005		
				(0.0006)		
EEA signals (normal)				-0.0010		
				(0.0007)		
border signals (wholesaler)					0.0005	
					(0.0008)	
border signals (normal)					-5.1×10^{-6}	
					(0.0013)	
history signals (wholesaler)						0.0002
						(0.0004)
history signals (normal)						0.0002
						(0.0006)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.124	0.124	0.124	0.124	0.124	0.124
Observations	322,879	322,879	322,879	322,879	322,879	322,879

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table 18: Wholesaler linkages - signal clarity

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals (wholesaler)	0.0323*					
	(0.0187)					
matching signals (normal)	0.0573**					
	(0.0288)					
non-matching signals (wholesaler)		0.0061				
		(0.0075)				
non-matching signals (normal)		0.0044				
		(0.0098)				
total signals (wholesaler)			0.0077			
			(0.0075)			
total signals (normal)			0.0066			
			(0.0098)			
EEA signals (wholesaler)				0.0009		
				(0.0083)		
EEA signals (normal)				-0.0235*		
				(0.0123)		
border signals (wholesaler)					0.0123	
					(0.0112)	
border signals (normal)					-0.0100	
					(0.0167)	
history signals (wholesaler)						0.0233***
						(0.0071)
history signals (normal)						0.0250**
						(0.0111)
Peer characteristics	yes	yes	yes	yes	yes	yes
Firm destination experience	yes	yes	yes	yes	yes	yes
Firm characteristics	yes	yes	yes	yes	yes	yes
firm FE	yes	yes	yes	yes	yes	yes
destination FE	yes	yes	yes	yes	yes	yes
year FE	yes	yes	yes	yes	yes	yes
nace4 FE	yes	yes	yes	yes	yes	yes
R^2	0.124	0.124	0.124	0.124	0.124	0.125
Observations	322,879	322,879	322,879	322,879	322,879	322,879

Clustered (firm) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*