Exporters and their Networks

The role of domestic network linkages in export entry

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

November 2025

Abstract

In this paper, we investigate whether firms' domestic network linkages can facilitate export participation. Using rich data on buyer-seller linkages in the Belgian production network, we find that network heterogeneity is a key determinant of the extensive margin of trade. Firms linked to experienced exporters via their business transaction network are more likely to access foreign markets but show no difference in export behavior upon entry. This suggests that network effects operate through a sunk cost channel and promote entry by reducing destination-specific costs associated with market access. We show that domestic linkages capture a new dimension of firm heterogeneity that facilitates entry alongside other firm characteristics. Determinants of export participation are thus generated inside and outside of the firm.

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

JEL Classification: F12, F14, D83, D85

*We thank Pol Antras, Paola Conconi, Ronald Davies, Xiang Ding, Sharat Ganapati, Kyle Handley, Keith Head, Colin Hottman, Beata Javorcik, Amit Khandelwal, Nuno Limao, Isabelle Méjean, Marc Melitz, Ferdinando Monte, Ralph Ossa, Charly Porcher, Christian Proebsting, Margit Reischer, Massimo Riccaboni, Rowan Shi, Gonzague Vannoorenberghe, Jaume Ventura and participants of the Leuven-Louvain trade workshop, NBB internal seminar, RIEF Doctoral Meeting, EOS conference, SMYE 2022, QMUL PhD Workshop, WIEN Workshop 2022 and lunch seminars at Georgetown and the University of Maryland for valuable comments and discussions. Philipp Ludwig gratefully acknowledges funding from the Research Foundation Flanders (FWO) and thanks the National Bank of Belgium for their hospitality during a research stay where this paper was developed.

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

[△]Lee Kuan Yew School of Public Policy, National University of Singapore. p.ludwig@nus.edu.sg

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

1 Introduction

Export participation remains a rare phenomenon with only 4-5% of firms directly engaging 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 (Eaton et al., 2009). Trying to understand why only a handful of firms can overcome entry barriers and access foreign markets therefore remains an important question.

Our paper contributes to this debate by investigating how heterogeneity in firm networks shapes export behavior. Traditionally, low export participation has been related to the presence of sunk entry costs, allowing only a select number of firms to be profitable in a foreign market. This insight sparked an influential literature of heterogeneous firm trade models which links export entry decisions to individual firm characteristics. Prominent attributes determining export participation are firm productivity (Melitz, 2003), access to finance (Manova, 2013), experience in similar markets (Albornoz et al., 2012; Morales et al., 2019), and economies of scope in the destination (Arkolakis et al., 2021).

What all these papers share is that they link export behavior to *internal* firm characteristics. In this paper, we introduce the idea that network linkages generate a set of *external* entry determinants that are based on the characteristics of buyers and sellers in the firm's domestic production network. Domestic networks not only coordinate the exchange of goods and services but also connect potential entrants to firms with direct export experience. If any export-related information diffuses along network linkages, export participation may not just depend on internal firm characteristics but also vary with the amount of export information provided by the network. As each firm network is unique, the amount and content of export information that firms are exposed to differs from firm to firm, even among firms that share equal levels of productivity. Network exposure therefore represents a distinct form of firm heterogeneity that is not captured by internal firm characteristics.

Our paper allows for this network dimension of firm heterogeneity and empirically investigates how internal and external entry determinants affect the export participation of Belgian firms. We start by developing a stylized model of export entry in which firms are heterogeneous in their characteristics and network linkages. Interactions with other members of the network, which we refer to as network peers[1] throughout the paper, generate idiosyncratic shocks that affect entry decisions by raising export profits or reducing the sunk cost of entry. These shocks capture a wide variety of non-pecuniary network externalities such as productivity spillovers or a diffusion of export-related information. The key novelty of the model is to use functional form assumptions that are inspired by the literature on social networks (Bramoullé et al., 2009) and spatial economics (Anselin et al., 2008) to express these shocks as functions of distinct network characteristics. This allows the model to separate network effects into different channels (productivity spillovers vs export market information), control for the specificity of the exchanged information (market-specific vs exporting in general), and capture a general form of diffusion that does not require firms to operate in close spatial proximity.[2] While the empirical framework establishes a direct relationship between export participation and network linkages, it remains fully agnostic as to whether network linkages truly facilitate entry and through which channel they operate. We exploit this flexibility to test and compare competing hypotheses when taking the model to the data.

From the model, we derive an estimation equation that takes the form of a time-space recursive model and relates entry decisions to the firm's own characteristics, as well as the characteristics and export experience of network peers. Empirically, the key mechanism we explore is to what extent *current* export experience in the network facilitates *future* access to foreign markets for connected firms. For this purpose, we exploit changes in the export behavior of network peers and define observed entry decisions to new export

[1]The term network peer in this paper describes any (direct) buyer or supplier interacting with the firm in the production network. While our main analysis focuses on interactions with buyers (backward linkages), we also consider linkages to suppliers in section 4.1.

[2]We thereby contribute to an existing spillover literature which has studied correlated import and export behavior of firms located in close geographic proximity (Koenig et al., 2010; Fernandes and Tang, 2014; Bisztray et al., 2018).

destinations as signals that carry valuable entry-related information. We then assess how network effects in the form of exposure to export signals and interactions with productive peers affect the probability of entering a particular foreign market in the next period.

To identify the impact of network effects on entry, we initially assume that the formation of network linkages and the export behavior of peers are exogenous. Later, we relax this assumption by introducing a network selection model as in Arduini et al. (2015) and Qu et al. (2017) to control for endogenous network formation and develop a network-based instrument to control for correlated export behavior that is driven by common shocks rather than information diffusion.

Using detailed balance sheet, trade, and business transaction data from the universe of Belgian firms, we are able to capture each firm's entire domestic production network and export behavior for the years 2002-2014. We use this data to estimate firm and network productivity, track the diffusion of export signals, and empirically estimate how network effects influence the export decisions of Belgian firms.

Bringing the model to the data yields four important results. First, we find that network effects are an important determinant of export participation. Even after controlling for internal firm characteristics like productivity, each incoming export signal on average increases the entry probability for a particular market in the next period by 0.43 percentage points. Signal effects are destination-specific and equivalent to a 13% increase in productivity of the signal-receiving firm. Conversely, interactions with productive network peers have no effect on entry rates to specific export destinations. These results suggest that export participation is not uniquely determined by internal firm characteristics, but also responds to external entry determinants in the form of export signals.

Second, we show that network effects facilitate foreign market access by reducing the sunk cost of entry. Receiving an export signal has no effect on export revenues, neither for export starters nor for incumbent exporters. This lack of response on the intensive margin of trade is inconsistent with interpretations that link network effects to changes in operating profits or variable cost. Instead, it strongly suggests that network effects operate exclusively through a sunk cost channel. Similarly, we cannot find any empirical support for an interpretation of export signals as devices that reduce demand uncertainty. Signals facilitate entry regardless of whether they indicate favorable or unfavorable market conditions and have no impact on post-entry survival rates in the foreign market. We therefore interpret signals as externalities that raise export participation by lowering the initial cost burden associated with market entry, such as destination-specific import regulations and procedures.

Third, we study how network effects shape export participation on aggregate. While the individual effect of incoming signals is large, aggregate network benefits are held back by a low prevalence of signal diffusion. Providing all firms with just one additional signal would increase the annual number of export starts by 3-5%. Policies that raise signal prevalence, for example by facilitating the formation of network linkages among domestic firms, could therefore lead to a substantial increase in export participation and connect more firms to global markets.

Finally, we find that marginal network effects exhibit decreasing returns to network scale. Receiving an identical signal from a larger network has a smaller effect on entry. We show that this result is closely related to negative assortative matching in the network formation process. The share of network linkages to exporters falls with the size of the network. This increasing exposure to non-exporters generates network noise which acts as an attenuating force for the beneficial impact of export signals. This finding highlights an important difference between internal and external determinants of export participation. While all firms benefit from higher levels of productivity, network effects appear to be particularly important for connecting small firms (with small networks) to foreign markets.

Related literature: By introducing network heterogeneity as a new determinant of the extensive margin of trade, our paper contributes to three broad strands of literature. A first strand studies production networks and explains how search costs, matching frictions, and two-sided firm heterogeneity influence the formation of buyer-seller linkages (Arkolakis et al., 2023; Bernard et al., 2018, 2022; Chaney, 2014; Dhyne et al., 2021; Eaton et al., 2022; Fontaine et al., 2023; Eaton et al., 2022, 2025; Huang et al., 2024; Panigrahi, 2021). We contribute to this literature by highlighting that diffusion in domestic networks facilitates export entry. As the majority of trade is in intermediates (Johnson and Noguera, 2012), entries often coincide with the creation of a cross-border linkage. This shows that domestic networks play an active role in the formation of international networks.

We also speak to a literature on trade intermediation which has emphasized the role of wholesalers in connecting domestic firms to foreign markets (Ahn et al., 2011; Bernard et al., 2010, 2018; Connell et al., 2019; Fujii et al., 2017; Ganapati, 2025). Our results show that business interactions with any firm, including non-wholesalers, can promote export participation, suggesting that information diffusion is a much broader mechanism than previously thought.

Finally, our paper is related to an extensive literature on peer effects in networks (Advani and Malde, 2018; Boucher et al., 2024; Bramoullé et al., 2020) which has developed flexible empirical frameworks that relate agent outcomes to network activity. Introducing this methodology to international trade not only underlines the framework's usefulness across different economic domains, but also illustrates a direct connection between the returns to network scale and the direction of assortativity in the network formation process. As this framework has applications across a wide variety of networks and fields, understanding the direction of assortativity should therefore become one of the first objects of interest in any empirical setting that relates network effects to agent outcomes.

This paper has 5 sections. Section 2 presents our empirical framework and discusses identification. Section 3 describes data sources and sample selection. Results are presented in section 4 and discussed in section 5.

2 Empirical framework

What does it take to become an exporter? Our goal is to provide a new perspective on this question by considering both firm characteristics and network interactions in the domestic production network as determinants of export participation. To do so, we need an empirical framework which relates export entry decisions to both dimensions of firm heterogeneity. 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[3] and has commonly not been considered as a direct determinant of export participation.[4]

As a baseline, we start with a stylized model of export entry in which firms initially only differ in productivity. We then introduce network heterogeneity by allowing operating profits and barriers to entry in foreign markets to respond to cost and information shocks that propagate along network linkages. As each firm has a unique set of linkages, exposure to network shocks generates additional variation in entry decisions above and beyond the firm's own characteristics. The network-augmented model borrows identification strategies from the social network literature, produces an intuitive estimation equation that can readily be taken to the data and allows us to explore distinct channels through which networks affect firm behavior on the extensive and intensive margin of trade.

2.1 Baseline entry framework

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

$$Pr(y_{id,t} = 1) = Pr\left(\frac{\bar{\Pi}_{id,t}}{r} > \bar{f}_{d,t}\right) \tag{1}$$

Suppressing time subscript t, firm profit in market d is $\bar{\Pi}_{id} = p_{id}q_{id} - a_i w_i q_{id}$ where p_{id}, q_{id}, w_i and a_i represent price, quantity, nominal wages and (inverse) firm productivity.[5] Under the standard setting of monopolistic competition and CES utility, demand for products of

[3] For an overview of the role of networks in international trade see Chaney (2016).

[4] Notable exceptions are found in the literature of trade intermediation where authors relate export entry decisions to interactions with wholesalers (Connell et al., 2019; Bernard et al., 2018) and in the literature of endogenous network formation where authors relate the size and productivity of domestic networks to the formation of cross-border linkages (Arkolakis et al., 2025; Eaton et al., 2022). However, none of these papers allows for networks to differ in export experience.

[5] Inverse productivity a_i is defined as the units of labor needed to produce one unit of q_i.

firm i 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 is $p_i = \frac{\sigma}{\sigma-1} a_i w_i$, a constant markup over marginal cost $a_i w_i$. The final price faced by foreign consumers is $p_{id} = p_i \tau_d$ where τ_d represents an ad-valorem iceberg-type trade cost incurred when shipping goods to market d. Plugging the resulting profit term $\bar{\Pi}_{id}$ into equation 1, we can express a firm's entry decision as

$$Pr(y_{id,t} = 1) = Pr\left(\left[\frac{\sigma}{\sigma-1} \frac{a_{i,t} w_{i,t} \tau_{d,t}}{P_{d,t}}\right]^{1-\sigma} \frac{\mu_{d,t} E_{d,t}}{\sigma r} - \bar{f}_{d,t} > 0\right) \tag{2}$$

Equation 2 illustrates that entry behavior is exclusively determined by heterogeneity in firm characteristics $w_{i,t}$ and $a_{i,t}$. This is emblematic of Melitz (2003) type trade models in which firm productivity explains the sorting of firms into exporters and non-exporters.[6] Firms with higher productivity (lower $a_{i,t}$) are more likely to start exporting. In what follows, we allow networks to alter this behavior.

2.2 Augmented entry framework with network interactions

We consider each firm's domestic production network as a source of idiosyncratic shocks that propagate along network linkages. There are two types of network shocks: A profit shifter ω and a sunk cost shifter ϕ.

Before discussing them in more detail, we introduce two important assumptions that illustrate how network shocks augment the baseline entry framework. First, we interpret these shocks as non-pecuniary externalities and assume each to be i.i.d. with unit mean and constant variance.[7] This allows networks to generate idiosyncratic variation in entry

[6]The Koenig (2009) model focuses on export starters, excluding firms that never export ($y_{id} = 0 \ \forall \ t$) or export continuously ($y_{id} = 1 \ \forall \ t$). By focusing on export profit, it also abstracts from domestic market participation as in Melitz (2003). Despite these simplifications, entry equation 2 captures key characteristics of Melitz-type models by linking entry to firm productivity and sunk entry cost.

[7]Treating network effects as externalities assumes that the formation of domestic buyer-seller linkages is not driven by a strategic motive to learn about foreign markets. Our framework therefore abstracts from models with network games (König et al., 2019) where optimal firm and network 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. While we do not allow linkages to form endogenously with the intent to learn, we do account for endogenous network formation that arises from unobserved shocks which simultaneously affect

behavior without changing baseline equation 1.[8] Second, we assume that shocks only affect entry behavior in the subsequent period. From a modeling perspective, this conservative approach reflects the fact that network effects likely take time to evolve.[9] Empirically, a lagged response addresses a data limitation that prevents us from identifying the sequence of agent and network decisions if both take place in the same year.[10] In the presence of network externalities, the *realized* export profit and sunk entry cost of firm i are thus defined as $\Pi_{id,t} = \bar{\Pi}_{id,t} \cdot \omega_{i,t-1}$ and $f_{id,t} = \bar{f}_{d,t} \cdot \phi_{id,t-1}$.

To relate shocks to each firm's individual set of network linkages, we make functional form assumptions which are inspired by the literature on social networks (Bramoullé et al., 2009; Calvó-Armengol et al., 2009). The key idea is to express profit and sunk cost shifters as functions of the distinct characteristics (x_j) and foreign market experience (y_{jd}) of all buyers and sellers j that firm i interacts with in the domestic production network. Following the literature, we define shocks as

$$\ln \omega_{i,t} = \sum_j \bar{s}_{ij,t} x_{j,t} \tag{3a}$$

$$\ln \phi_{id,t} = \sum_j s_{ij,t} y_{jd,t} \tag{3b}$$

where $x_{j,t}$ represents time-varying characteristics of network peers, $y_{jd,t}$ is an indicator variable[11] for export starts of network peers, and $s_{ij,t}$ ($\bar{s}_{ij,t}$) are elements of a (row-normalized) binary interaction matrix S_t[12] which captures all domestic firm-to-firm interactions in the economy in year t.

This characterization of shocks captures two distinct channels through which network

network formation and export entry in section 2.4.

[8] Because shocks enter the model multiplicatively, assumptions for ω and ϕ ensure that the present value of future export profits, optimal demand, and mill prices remain unchanged. It implies that firms are oblivious to any potential benefits provided by their network.

[9] Cost shocks, for example, are unlikely to affect within-period profits if existing contracts introduce stickiness to input use. Information shocks may not immediately reduce entry barriers if processing and acting on the new information takes time.

[10] Empirical network proxies only vary at the year level. Within a given year, we therefore cannot observe who acted first: firm i or its network.

[11] We refrain from weighting export signals due to a lack of a theory-consistent weighting scheme. Instead, we explore signal heterogeneity empirically in section D.2 by studying how network effects differ across a range of peer and linkage characteristics.

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

externalities affect export participation. First, it allows export profits to vary with the average characteristics of network peers x_j. This channel encompasses mechanisms such as Hicks-neutral productivity spillovers, quantity discounts from pooled procurement, or freight savings from consolidated shipments.[13]

Second, we allow sunk cost f_{id} to vary with the number of export starts y_{jd} that occur in firm i's domestic production network. An export start is defined as the first entry of a firm into the foreign market.[14] We interpret export starts as events that generate new information about foreign markets. Each time a firm j in the network of firm i enters a new destination d for the very first time, it pays sunk cost f_d to assess local demand preferences, identify local distribution networks, and learn how to comply with local standards. A diffusion of this information can therefore promote entry by reducing the sunk cost that *other* firms need to incur to enter the same market. Channel 2 captures this idea by allowing destination-specific information to diffuse along network linkages s_{ij} to connected firms which treat incoming information as an *export signal*. In this paper, an export signal therefore refers to an export entry event that occurs in firm i's domestic production network and we interpret signals as destination-specific cost shocks for sunk cost f_{id}.

The introduction of network externalities into the baseline framework can explain why firms with identical characteristics make different entry decisions. Profit shifter ω_i is determined by the characteristics of network peers (x_j) and generates variation in entry behavior at the *firm-level*. Two identical firms (x_i) that interact with different network peers (x_j) may not reach the same export decision if one peer group is more productive than the other. Sunk cost shifter $\phi_{id,t}$ goes even further by allowing entry barriers to vary with the level of export experience (y_{jd}) in the network. This generates entry variation at the *firm-destination level*. Two identical firms (x_i) that interact with equally productive

[13]For a recent paper studying productivity spillovers using similar functional form assumptions, see Iyoha (2023).

[14]Empirically, we consider cross-border transactions as export starts if the firm has not served the foreign market for at least two consecutive periods. This definition applies to all entry decisions taken by firm i or network peers j. It implies that no firm can enter the same market in two consecutive periods.

peer groups (x_j) can still take different export decisions because they face different entry barriers across export destinations. Export participation is thus no longer exclusively determined by heterogeneity in the characteristics of firm i, but also linked to the characteristics and experience of buyers and sellers j that firm i interacts with in the domestic production network.

2.3 Advantages of the augmented entry framework

Our augmented entry framework provides a new perspective on how networks shape foreign market access but also introduces new identification challenges. Before discussing our identification strategy in section 2.4, we want to highlight several key advantages of our modeling approach.

Leveraging network experience: Since intermediate goods account for the majority of trade (Johnson and Noguera, 2012), an export entry will often coincide with the formation of a business linkage to a foreign firm. This connects our entry framework to the network formation literature which also studies the creation of cross-border linkages via search and matching models.[15] A common feature across these network formation models is that the success of a foreign match depends on the characteristics and the number of linkages of each individual firm. Firms exert effort to identify foreign clients and benefit from larger networks in the form of lower search costs. Although formation models and our augmented framework both allow firm and network heterogeneity to shape the formation of cross-border linkages, network effects operate at different levels of granularity. Search and matching models only consider heterogeneity in *network size* which implies that network effects differ across firms but not across destinations. Sunk cost shifters in our augmented framework instead account for variation in *export experience within networks*. This new channel can rationalize cases in which identical firms with equally sized networks choose to form linkages with firms in different export markets. Our empirical approach therefore introduces a new mechanism through which networks shape export behavior that has not been considered in existing search and matching models.

[15] Recent examples include Eaton et al. (2022), Arkolakis et al. (2025) and Eaton et al. (2025).

Wider network reach: By allowing profit and sunk cost shifters to propagate along all domestic buyer-seller linkages, network externalities can reach *any* firm in the entire economy.[16] The scope of our mechanism is therefore considerably wider than other processes such as agglomeration effects of nearby exporters, manager experience, human capital movements between firms or learning through trade intermediaries, which all have been identified in the literature as important avenues towards foreign market access.[17] Our approach complements these channels by casting a wider net and accounting for firms that may not benefit from agglomeration, hiring or indirect exporting.[18]

Accounting for network noise: Another advantage of our augmented framework is the ability to flexibly nest different modes of diffusion of sunk cost shifter ϕ into the entry framework. To illustrate this point, we define

$$\begin{aligned}
\text{signal intensity} &= \sum_j s_{ij,t} y_{jd,t} \\
\text{signal clarity} &= \sum_j \bar{s}_{ij,t} y_{jd,t} = \frac{\text{signal intensity}}{\sum_j s_{ij,t}} \\
\text{network noise} &= 1 - \text{signal clarity}
\end{aligned}$$

Signal intensity counts the number of incoming export signals in any given period. It thus captures the *absolute* amount of information a firm receives. Signal clarity instead is defined as the share of linkages that provide export signals. This *relative* measure accounts for the fact that networks differ in size and relates the amount of information to the size of the network it originates from. Distinguishing between these two measures becomes important if interactions that do not yield any signals exacerbate the uptake of incoming signals. In our context, this is likely the case because firms need to invest

[16]Firms that do not source from or sell to any other firm are dropped from the sample as explained in section 3.

[17]Agglomeration effects typically require agents to reside in close proximity to one another, which restricts the geographic reach of network effects to specific neighborhoods (Koenig et al., 2010; Fernandes and Tang, 2014; Bisztray et al., 2018). Benefits from hiring managers with foreign market expertise as in Mion et al. (2025) and Patault and Lenoir (2024) do not extend to firms whose workforce remains unchanged. Learning about product appeal in foreign markets by exporting indirectly through wholesalers as in Ahn et al. (2011) or Connell et al. (2019) ignores a potentially much larger benefit that stems from interactions with exporters in general.

[18]We account for these alternative channels empirically in section 4.1.

time and resources to engage with their network peers, irrespective of whether they learn something in the process. In that case, network noise, defined as the share of linkages that do not provide signals, acts as an attenuating force for network externalities. We build on this insight in section 4.4 to understand diverging network effects for firms with small and large networks.

Uncover underlying mechanisms: Profit and sunk cost shifters in the augmented framework capture a wide range of mechanisms through which networks affect entry behavior. By exploiting variation within and across shifters, we can infer which mechanism is operating in practice. Export signals y_{jd} for example reflect export experience for a particular export destination d. If entry barriers differ across foreign markets, a signal should only reduce the sunk entry cost for the unique market it originates from. To test this hypothesis, we distinguish between matching and non-matching signals and explore the relevance of each signal type empirically in section 4. The presence of shocks on both sides of entry equation 1, on the other hand, allows us to investigate whether network externalities only matter for market entry or for exporting in general. If profit shifter ω_i is an important determinant of operating profit Π_{id}, we would expect to see a significant impact of network externalities on both the intensive and extensive margin of trade. In that case, export starters and incumbents both benefit from lower variable cost or higher productivity. If network externalities instead only affect entry, they must be operating through a sunk cost channel as pointed out by Chaney (2008). By applying our framework to both the intensive and extensive margin of trade, we can thus learn more about the underlying mechanism of network effects.

Pathways to identification: To derive our empirical estimation equation, we introduce profit and sunk cost shifters into entry equation 1 and express network externalities via equations 3a and 3b. Denoting vectors in bold and scalars in plain typeface, our network-augmented entry equation in logs becomes

$$Pr(y_{id,t}=1) = Pr\left(\boldsymbol{\gamma}'\boldsymbol{x_{id,t}} + \boldsymbol{\delta}'\sum_j \bar{s}_{ij,t-1}\boldsymbol{x_{j,t-1}} + \beta\sum_j s_{ij,t-1}y_{jd,t-1} + \alpha_{d,t} + \varepsilon_{id,t} > 0\right) \quad (5)$$

where vectors $\boldsymbol{x_{id,t}}$ and $\boldsymbol{x_{j,t}}$ collect seller and buyer characteristics related to export entry, $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ and β are parameter (vectors) to be estimated, $\alpha_{d,t}$ collects firm-invariant constants, and $\varepsilon_{id,t}$

is an idiosyncratic error term.[19]

By relating outcome $y_{id,t}$ to firm i's own characteristics $x_{id,t}$ and to lagged characteristics $(x_{j,t-1})$ and outcomes $(y_{jd,t-1})$ of it's network, our equation closely resembles models in the peer effects (Manski, 1993; Bramoullé et al., 2009) and spatial economics (Anselin et al., 2008; Qu and Lee, 2015) literature.[20] This represents another key advantage of our framework, as both strands have established strong econometric foundations that guide the identification strategy presented below.

2.4 Identification

We start by rewriting entry equation 5 in matrix notation. Firm characteristics, linkages, and export starts in a given year are collected in matrices X, S and Y, respectively. Using superscripts to indicate the underlying source of variation, we get

$$Pr\left(Y_t^{(id)} = 1\right) = Pr\left(\boldsymbol{\gamma}' X_t^{(i)} + \boldsymbol{\delta}' \bar{S}_{t-1} X_{t-1}^{(j)} + \beta S_{t-1} Y_{t-1}^{(jd)} + \alpha_{d,t} + \varepsilon_t > 0\right) \tag{6}$$

Our goal is to obtain estimates of parameters $\boldsymbol{\delta}$ and β which capture the causal effect of network externalities on export entry.[21] This requires network characteristics, linkages, and export experience to be uncorrelated with contemporaneous outcome error ε_t. As a benchmark, we impose sequential exogeneity by assuming $E\left(\varepsilon_t | S_{k-1}, X_{k-1}^{(j)}, Y_{k-1}^{(jd)}\right)$ $\forall\, k \leq t$ which implies that networks form exogenously after conditioning on observable firm and network characteristics. This rules out settings in which forward looking firms create linkages strategically in anticipation of future network benefits. In our context of export entry, strategic behavior implies that firms can keep track of all entry-related peer characteristics of their entire network and form precise predictions about *when and where* their peers are expected to enter a foreign market. This assumption does

[19] $\boldsymbol{x_{id,t}} = \ln(a_{i,t} w_{i,t})^{1-\sigma}$ and $\alpha_{d,t} = \ln((\sigma/(\sigma-1))^{1-\sigma} P_{d,t}^{\sigma-1} \mu_{d,t} E_{d,t} (\sigma r)^{-1})$ capture firm-specific characteristics and firm-invariant constants of expected profit $\bar{\Pi}_{id,t}$, respectively. Network linkages $s_{ij,t}$, characteristics $\boldsymbol{x_{j,t}}$ and export starts $y_{jd,t}$ are formally defined in section 3.

[20] In settings with contemporaneous network effects, the peer effects and spatial economics literature refers to these as local aggregate and spatial autoregressive models, respectively. Closest to our specification with lagged network effects are time-space recursive models in spatial economics (Halleck Vega and Elhorst, 2017), where time-space refers to temporal and spatial lags in the network components. A key difference to the literature is that we do not need to consider lagged outcomes of firm i as an additional control. This type of autocorrelation cannot occur in our setting, due to our definition of an export starts. As explained further in section 3, we only consider exports to destination d as an export start if the firm has not served that market for two consecutive periods.

[21] We use a signal intensity specification $(\beta S_{t-1} Y_{t-1}^{(jd)})$ to discuss identification challenges. To move to a signal clarity specification, simply replace binary interaction matrix S with row-normalized interaction matrix $\bar{S}$. All arguments made below hold for both specifications.

not hold in practice. Endogeneity linked to a strategic anticipation of network effects is therefore unlikely to render profit and sunk cost shifters endogenous.

A more tangible concern in this benchmark setting are unobserved shocks, which pose three main identification challenges: First, they create a simultaneity problem that threatens the separate identification of network effects $\boldsymbol{\delta}$ and β. Second, they can create correlation between firm and network outcomes that renders $Y_{t-1}^{(jd)}$ endogenous and results in a biased estimate of network effects. Finally, they can jointly affect the formation of domestic (production network) and cross-border (exporting) linkages which renders S_{ij} endogenous and introduces a selection bias to our estimates. In line with the social network and spatial economics literature, we refer to these challenges as a *reflection problem*, *correlated effects* and *endogenous network linkages* accordingly.[22] We discuss each of these challenges sequentially to dissect one issue at a time.

The reflection problem: To study how profit and sunk cost shifters affect export participation, we need to ensure that network parameters $\boldsymbol{\delta}$ and β can be identified separately. This requires a sufficient overlap across firm networks.[23] If interactions are limited to other members of the same network, $\boldsymbol{\delta}$ and β become perfectly collinear because in the presence of common shocks all firms within the same cluster act simultaneously and there is no variation from cross-cluster linkages to separately determine the impact of network characteristics $X^{(j)}$ and export experience $Y^{(jd)}$ on outcome $Y^{(id)}$.

Bramoullé et al. (2009) and Liu et al. (2014) show how to overcome this *reflection problem* in local-average and local-aggregate models which we refer to as signal clarity and intensity as explained above. In local-average models, network effects $\boldsymbol{\delta}$ and β are identified separately if identity matrix I and interaction matrices S and S^2 are linearly independent. In local-aggregate models separate identification requires the row sums 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 uni-directional, which ensures linear independence of network matrices due to the presence of intransitive triads[24] and the fact that each seller interacts with

[22]Our identification strategy builds on previous work in social networks and spatial economics (Bramoullé et al., 2020; Advani and Malde, 2018) and therefore adopts terminology common to that literature. All *italic terms* are explained in detail in the following sections.

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

[24]An intransitive triad describes a network structure where firm A interacts with firm B, B interacts with firm C, but C does not interact with A.

a different number of buyers, leading to a non-constant row sum of S.

Correlated effects: While overlapping networks allow us to identify $\boldsymbol{\delta}$ and β separately, we still need to ensure that these parameters capture the causal effect of network externalities, rather than a correlation in firm and network behavior that is driven by common shocks. These *correlated effects* are a concern in our setting because buyers and sellers naturally face various domestic and foreign shocks that can alter their export participation decision, but that remain unobserved by the econometrician.

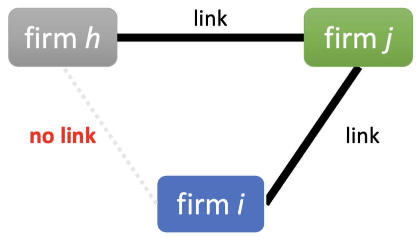
Many of these unobserved shocks can be accounted for via high-dimensional fixed effects (FE). Common examples are foreign demand shocks (destination-year FE) or foreign market appeal (firm-destination FE). A unique feature of our setting is that we can extend this approach to control for unobserved variation within firms over time. By adding firm-year FE to equation 6, we can control for all conventional sources of firm heterogeneity and still recover estimates of parameter β. This allows us to separate the effect of network experience from competing channels, such as local agglomeration effects, firm productivity, and productivity spillovers from network peers.[25]

But not all correlated effects can be addressed via fixed effects. An important example in our setting are common shocks that introduce serial correlation to outcome error ε. To illustrate this point, consider a sudden surge in popularity of Belgian chocolates in the Chinese market. While some Belgian chocolate producers will immediately seize the opportunity, others will require more time to respond, as ramping up production and establishing supply chains for their precious cargo are non-trivial tasks. A delayed response to common shocks exacerbates the identification of network effects because the shock-induced variation in firm and network behavior is observationally equivalent to the network externality we are trying to capture.

This form of correlated effect is a particular concern for network signals Y_{t-1}^{jd}. As we cannot control for common shocks at the firm-destination-year level via fixed effects, we have to treat export signals as endogenous and find a suitable instrument to obtain an unbiased estimate of β.

[25] As firm-year FE absorb many variables of interest from the stylized entry model, including the effect of profit shifter $\boldsymbol{\delta}$, we treat firm-year FE specifications as a robustness check and do not include them in our benchmark equation. Results with firm-year FE are presented in section D.2.

Figure 1: Instrumentation strategy



Our solution is an instrument that exploits temporal and spatial lags in the network structure: export starts of second-order peers $\left(Y_{t-2}^{(hd)}\right)$. Second-order peers h are firms that directly interact with firm j but have no observed link to firm i as illustrated in Figure 1. By virtue of the augmented framework, export starts of second order peers in period $t-2$ act as sunk cost shifters for entry decisions of firm j in period $t-1$ and are thus directly correlated with endogenous regressor Y_{t-1}^{jd}. Moreover, we argue that our instrument is plausibly orthogonal to outcome error ε_t. First, actions of firm h take place *before* the realization of common shocks in $t-1$. Second, they are performed by firms that are spatially distant from affected firms j and i. This spatial lag alleviates concerns in settings where a common shock is persistent and thus potentially correlated with firm behavior in $t-2$. This is due to the fact that a destination-specific shock that affects all three groups of firms h, j, and i is likely large enough to be sufficiently absorbed by destination-year FE.[26]

The validity of this instrument relies on the assumption that second-order peers h are only linked to firm i via intermediate firm j. To corroborate the credibility of this exclusion restriction, we exploit the full network structure[27] and exclude all firms from the set of second-order peers that have a first-, third-, fourth- or fifth-order linkage to firm i. To curb the influence of unobserved firm interactions that occur outside the production network, we only consider firms h that are located outside of the province of firm i. We address any remaining concerns by clustering standard errors at the firm-level to allow for arbitrary serial correlation within firms over time and

[26]Recall that our main concern are unobserved shocks at the firm-destination-year level. If export decisions of firms i, j, and h are all affected by the same shock, most of the unobserved variation of the shock will be at the destination-year level which we already account for with fixed effects.

[27]A natural concern in this setting is that our network sample does not accurately capture all relevant linkages of each seller. While this is likely the case in practice, we expect our approach to perform reasonably well as most social networks are extremely sparse. Missing or misspecified network linkages should therefore only represent a small fraction of total linkages when compared to the correctly identified absence of linkages between most firms.

present results of our instrumentation strategy in section 4.1.

Endogenous network linkages: Finally, we turn to network linkages S. Although firms in our setting are unlikely to form linkages strategically in anticipation of network externalities, this does not guarantee that linkages are exogenous. This is because a firm's ability to identify suitable buyers or suppliers domestically might be systematically correlated with their likelihood of establishing linkages abroad in the form of exporting. Unobserved shocks, which affect both domestic and foreign link formation, are problematic because they render network linkages endogenous and introduce bias to estimated network effects.

To address this concern, we introduce a network selection model developed by Arduini et al. (2015) and Qu et al. (2017) that is described in detail in appendix A. The model allows us to express network endogeneity as an unobserved shock to domestic production network S and export decisions $Y^{(id)}$. It accounts for any selection bias that is caused by correlated linkage formation and entry decisions via a selection correction term $\hat{\Xi}$ that takes the familiar form of a Heckman-type mills ratio.[28] By adding $\hat{\Xi}$ as an additional regressor to equation 6, we can purge outcome error ε of any unwanted correlation with the formation process and recover unbiased estimates of network parameters $\boldsymbol{\delta}$ and β.

3 Data, empirical setting and descriptive statistics

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.

3.1 Data sources and sample selection

At the center of our analysis are three administrative datasets that 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

[28]Our approach introduces network endogeneity in form of a correlation between network formation and export entry errors $\xi_{ij,t}$ and $\varepsilon_{id,t}$. We believe this modeling choice is appropriate given that both processes involve a search for business partners that could be affected by common unobserved shocks. For an empirical application of this approach in the context of productivity spillovers see Iyoha (2023). Alternative modeling approaches, which link outcome errors to unobserved variables in the formation process (Goldsmith-Pinkham and Imbens, 2013; Hsieh and Lee, 2016), instead rely on Bayesian methods. To keep the estimation parsimonious, we abstract from these alternatives.

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 these 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.[29] 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 €.[30] 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 thus 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 that 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.[31] 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

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

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

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

output, exports, and imports.[32]

At the destination level, we only consider export markets that are outside the European Economic Area (EEA), as information frictions are expected to represent a much larger barrier to entry in those markets compared to highly integrated EEA countries.[33] Non-EEA destinations on average 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.[34] This ensures that sufficient time has passed for market conditions to change such that information costs again become a relevant barrier to entry.[35] This implies that all observations of the first two years are dropped from our sample, reducing the sample time frame to 2004-2014. Further, we only consider firm-destination pairs with at least one export start across years to facilitate comparisons across different estimation approaches.[36]

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

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

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

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

[35]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.

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

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. Buyers have a clear incentive to communicate and enforce product requirements and standards when sourcing inputs from their suppliers. In fact, empirical evidence suggests that buyers in domestic production networks regularly perform audits and control visits, and provide technical assistance and trainings for their suppliers, which all contribute towards a diffusion of information along backward linkages.[37] Forward linkages are less likely to serve a similar purpose because sellers have little incentive to communicate export-related information to their buyers after a transaction is completed. We therefore focus on backward linkages as the main direction of diffusion.[38]

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 export signals. Suppliers which only account for a small share of total buyer sourcing may receive no information because the small transaction size does not necessitate any communication with buyers or indicates a low level of importance attached to the sourced input. We therefore need to distinguish between relevant and non-relevant network linkages and exclude those which are too small to play any meaningful role for the diffusion of export signals. To do so we compute the share 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 K_t. An interaction is defined as relevant for diffusion if suppliers

[37]Survey evidence by Alfaro-Ureña et al. (2022), who study interactions between multinational buyers and domestic suppliers in Costa Rica, suggests that 69% of buyers claim to provide, and 44% of suppliers acknowledge receiving direct support along backward linkages.

[38]An exception to this rule are wholesalers and freight-forwarders that provide highly specialized services such as warehousing, transport and logistics. These services are highly destination-specific and thus might contribute to a diffusion of export information from sellers to buyers. To explore this channel, we study the impact of network diffusion along forward linkages in appendix D.2.7. In line with our prediction, there is no beneficial impact of export signals on entry once we remove wholesalers from the set of sellers. Information diffusion along forward linkages thus appears to be a much more narrow process than the far-reaching benefits provided by signal diffusion along backward linkages as presented in section 4.

account for at least 1% of buyer sourcing. Interactions that account for less than 1% of buyer sourcing[39] are treated as irrelevant for information diffusion and are excluded from the sample.[40] Applying this rule to all entries of transaction value matrix K_t leads to a binary interaction matrix S_t[41] 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 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 62,000 firms, 25,000 export starts to 188 non-EEA destinations, and more than 1,000,000 firm-to-firm interactions per year between 2004 and 2014. The combined data allows us to trace the diffusion of export signals along network linkages and relate it to the export entry behavior of Belgian firms. To understand how each data source contributes to the analysis, we present descriptive evidence to illustrate firm behavior at the extensive margin of trade, the prevalence of export signals, and how the size of a network affects the diffusion process.

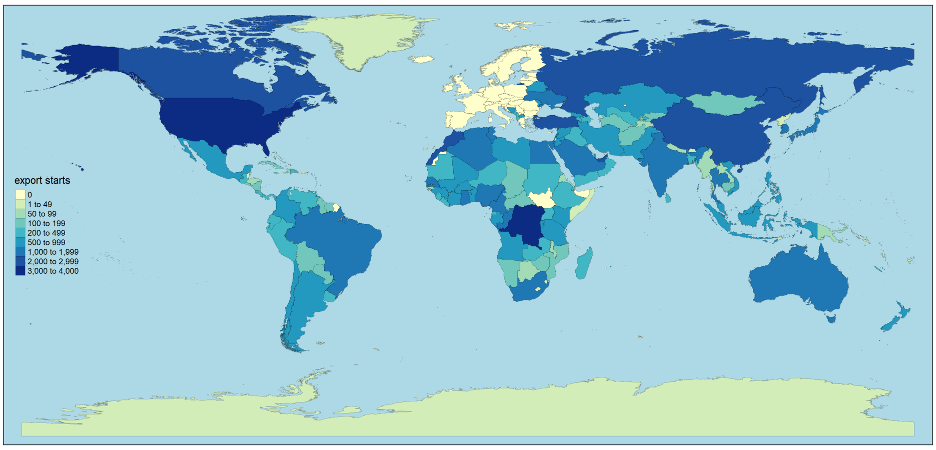
Extensive margin of trade: Figure 2 shows the spatial distribution of non-EEA export starts of Belgian firms between 2004-2014. Export decisions follow the rules of gravity and mainly occur in markets that are attractive due to their large size or proximity to Belgium. One exception is the concentration of new export activity 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 pattern shown in appendix B.1 is the distribution of entry decisions across geographic regions. While large countries like the US individually account for the largest number of export starts, more than two-thirds of non-EEA entries occur in Africa

[39] Our empirical results are robust to alternative thresholds as demonstrated in section 4.1.

[40] 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 make up 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 overall sample size.

[41] Our baseline model does not differentiate between transactions beyond the 1% threshold. To learn more about the role of interaction strength for network externalities, please see section 4.1.

Figure 2: Geographic distribution of export starts (2004-2014)



and Asia. As these regions contain a large number of countries with distinct import regulations, consumer preferences, and local supply networks, Belgian firms face considerable entry barriers when trying to access these markets. To reach the large consumer base in these emerging markets, network externalities might play an important role in facilitating market access.

Prevalence of export signals: The data allows us to identify over 728,000 export signals received by sellers between 2004-2014. We distinguish between matching and non-matching export signals to indicate whether the origin of an incoming signal matches the destination of a seller's subsequent export start. Matching signals therefore represent information of direct relevance to foreign market access whereas non-matching signals capture the general availability of export-related information in the network.

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
Number of firms: 61,685				

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. *Any signal* here refers to the sum of matching and non-matching signals.

Table 1 illustrates the prevalence of both signal types in each year and over the entire sample period. Despite the large number of signals diffusing along network linkages, receiving an export signal remains a rare event. Each year only 5.5% of sellers (around 3,390 firms) benefit from matching export signals which means that many entry decisions are still taken in absence of network externalities. This implies that there remains a large amount of cross-sectional variation we can exploit for our empirical analysis.[42] The distribution of firms receiving export signals also appears to be highly 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.[43]

Network externalities: The concentration of signals among a small number of sellers is directly related to the underlying linkage distribution. While seller networks on average consist of 14 different buyers including 2 exporters and 1 export starter, the overall distribution of seller linkages is highly skewed.[44] Figure 10 in appendix B.3 shows that while 5% of sellers only interact with a single buyer, sellers in the top decile on average interact with over 100 buyers in 2014. These vast differences in network size are closely related to seller characteristics. As shown in Figure 12 in appendix B.5, sellers with higher productivity on average interact with more buyers - a common pattern in production networks (Bernard and Zi, 2022; Bernard et al., 2022). Productive sellers can offer products at higher quality or lower prices and thereby attract a larger number of buyers.

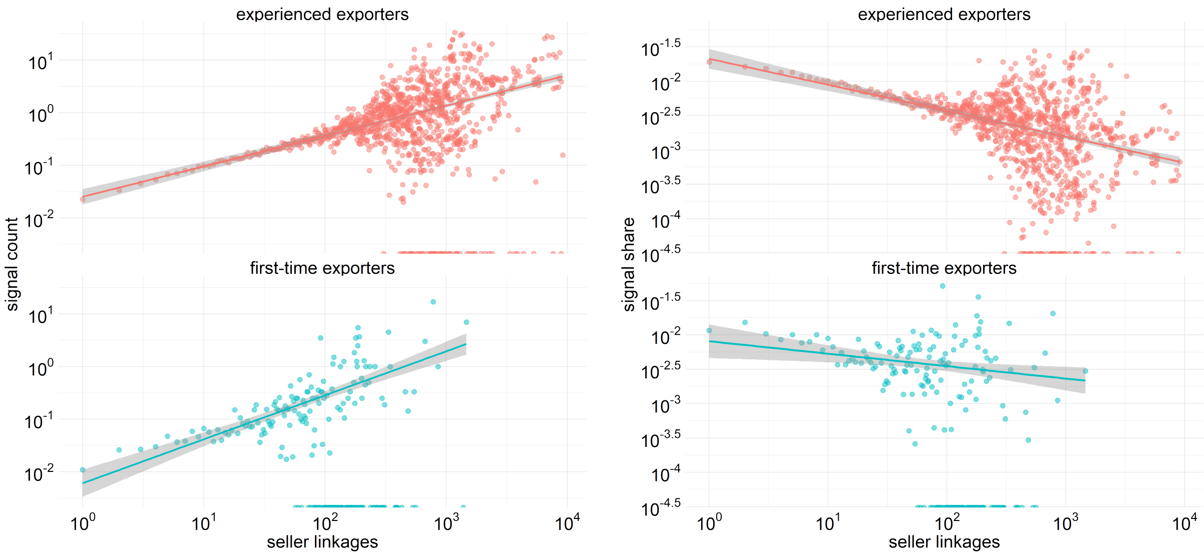
In our setting, this positive correlation between seller productivity and network size determines the amount of information each seller receives. Productive sellers with large networks mechanically receive a larger number of export signals as each additional linkage raises the probability of signal diffusion. Signal intensity is therefore increasing in network size as shown in Figure 3a. This advantage of large networks, however, can be misleading. Figure 13 in appendix B.5 shows that the share of exporters in seller networks is decreasing with network size. Highly productive sellers thus interact with a larger number of buyers, but on average are less likely to encounter an exporter. As a result, signal clarity, defined as the amount of information per linkage, decreases with network size as seen in Figure 3b. Firms with larger networks therefore experience higher

[42] If most sellers received export signals in every period, identification of network externalities would have to rely on within-firm variation in incoming export signals over time. Table 1 shows that our analysis can exploit both within- and between-firm variation when estimating network effects.

[43] The full distribution of export signals is shown in appendix B.2.

[44] Seller linkages to exporters and export starters are equally skewed as seen in appendix B.3.

Figure 3: Network externalities and network size



(a) signal intensity

(b) signal clarity

The graphs depicts empirical measures of network externalities which are defined in section 2.3.

signal intensity but lower signal clarity. This indicates that network effects may not linearly increase in network size, something we explore empirically in section 4.4.

3.3 Estimation

We now present our empirical framework. To bring entry equation 6 to the data, we need to make a distributional assumption for outcome error ε_t. As a starting point, our benchmark estimation uses a linear probability model with fixed effects (LPM-FE) which assumes that errors ε_t are *i.i.d.* and follow a normal distribution. A key advantage of the LPM-FE is the ability to flexibly accommodate high-dimensional fixed effects which allows us to control for unobserved shocks in domestic and foreign markets that may otherwise give rise to correlated effects as discussed in section 2.4. At the same time, the 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 the LPM-FE model, we 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 to the LPM-FE via average partial effects.[45]

Under normally distributed errors, we estimate the following reduced-form equation based on our time-space recursive lag model:

$$Pr\left(Y_t^{(id)} = 1\right) = Pr\left(\gamma' X_t^{(id)} + \delta' \bar{S}_{t-1} X_{t-1}^{(j)} + \beta S_{t-1} Y_{t-1}^{(jd)} + \psi_i + \psi_{d,t} > \varepsilon_t\right) \tag{7}$$

Seller export starts $Y_t^{(id)}$ are related to their own characteristics $X_t^{(id)}$, network effects in the form of buyer characteristics $X_{t-1}^{(j)}$ and export signals $Y_{t-1}^{(jd)}$ and a set of fixed effects ψ. We summarize the variables contained in each component below and present additional details in appendix C.2.

i. Seller characteristics $X_t^{(id)}$ capture determinants that affect seller export decisions in the 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). Complementary to 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

[45]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.

destinations or destinations with historic ties.[46] Lastly, we control for a seller's overall export expertise via the share of export sales in total sales.

ii. Buyer characteristics $X_{t-1}^{(j)}$ capture network externalities that affect entry decisions across all export destinations. Through the lens of the augmented framework, they represent profit shifters that raise export profit. We include buyer sales and TFP to control for externalities in the form of economies of scale from increased sourcing, and productivity spillovers caused by interactions with productive buyers.

iii. Lastly, we employ two distinct fixed-effect (FE) specifications to control for correlated effects. In the benchmark case, we include firm and destination-year FE ψ_i and $\psi_{d,t}$. This allows us to control for unobserved differences in firm performance and time-varying demand shocks in foreign markets. A second and more stringent specification extends this to firm-year FE. In this case, fixed effects absorb *any* time-varying characteristic at the firm-level, including many variables of the baseline entry framework and network effects from buyer characteristics. To remain as close as possible to the theoretical framework, we therefore rely on the weaker FE specification for our benchmark estimates and present results under more stringent specifications as robustness checks.

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 which includes 22,133 unique firms. All variables have been aggregated to the firm-year level to facilitate interpretation. The reported number of observations therefore differs from the regression tables which capture entry decisions at the firm-destination-year level.

[46] 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)

4 Results

We now explore how network interactions affect the export behavior of Belgian firms. We first focus on the extensive margin of trade. Bringing our augmented entry framework to the data shows that networks are an important determinant of export participation, even after controlling for firm productivity. This result is robust to alternative estimation strategies and sampling choices, and it holds when addressing endogeneity via network-based instruments and a dyadic network selection model. In contrast, network externalities have no impact on the intensive margin of trade. We interpret these findings as evidence that networks reduce entry barriers to foreign markets rather than boosting export profits, and present empirical evidence to rule out alternative channels such as a reduction in demand uncertainty. We then turn to the aggregate impact of network effects. Under the current level of diffusion, the aggregate number of network-induced entries remains small because only a handful of firms receive export signals in any given period. Raising the prevalence of signal diffusion, however, leads to a large increase in export participation, making this an interesting area for policy interventions. Finally, we study how marginal network effects are related to firm and network size to understand how networks contribute to the observed firm-size concentration at the extensive margin of trade. We find that marginal network benefits decrease with network size and explain how this result is linked to negative assortative matching in the underlying formation process.

4.1 Extensive margin

Our empirical framework allows network characteristics and experience to alter export decisions of Belgian firms. To assess the empirical relevance of these new channels, we initially treat network linkages and signals as exogenous and estimate entry equation 7 via a LPM-FE. We then subject our benchmark estimates to a series of robustness checks to explore how model specification, sample selection, and endogeneity affect network effects. A detailed account of this process is presented in appendix D.2. In all specifications, we control for firm productivity and focus on network parameters $\boldsymbol{\delta}$ and β which capture the impact of network-induced productivity spillovers and export signals.

Matching vs. non-matching signals: Productivity spillovers operate at the firm-level and thus affect entry decisions across all export destinations. Export signals, however, carry market-specific information and are therefore expected to exclusively reduce entry barriers in one par-

Table 3: Benchmark results - signal type

Signal type:	matching	non-matching	total	EEA	border	history
Variables						
export signal	0.0043***	0.0000	0.0000	0.0000	0.0003	0.0002***
	(0.0011)	(0.0001)	(0.0001)	(0.0001)	(0.0004)	(0.0001)
peer TFP	0.0004	0.0003	0.0004	0.0003	0.0003	0.0004
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0027)	(0.0027)
Peer characteristics	✓	✓	✓	✓	✓	✓
Destination experience	✓	✓	✓	✓	✓	✓
Firm characteristics	✓	✓	✓	✓	✓	✓
firm FE	✓	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓	✓
R^2	0.0926	0.0926	0.0926	0.0926	0.0926	0.0926
Observations	469,770	469,770	469,770	469,770	469,770	469,770

This table shows how export signals and productivity spillovers affect the export decisions of Belgian firms. Results are based on equation 7 and estimated via a LPM-FE. Each column shows the marginal effect of receiving a different type of export signal on a seller's probability to start exporting. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

ticular destination. A natural way to explore this idea is to sort incoming export signals by destination and test whether signals only facilitate access to the market they originate from (matching signal) or can also promote entry in other markets (non-matching signal).

Table 3 summarizes our benchmark estimates of network effects. The full table is available in appendix D.1. Across specifications, spillovers linked to the characteristics of network peers ($\boldsymbol{\delta}$) have no significant impact on entry decisions of connected firms. Interactions with more productive buyers therefore do not appear to facilitate foreign market access. In contrast, firms are more likely to enter a foreign market if they receive a matching export signal (column 1). On average, each incoming matching signal increases the probability to start exporting to the same destination market by 0.43 percentage points which is equivalent to a 13% increase in productivity of the signal-receiving firm.[47] As we explicitly control for productivity differences across firms and firm networks, this result illustrates that heterogeneity in export experience *within* networks is an important determinant of export participation. Two equally productive firms interacting with equally productive networks can still choose to enter different markets because they do not have access to the same amount of information and thus face different entry barriers across destinations.

[47]For this comparison we take the estimated coefficient of log TFP from Table D.1. The equivalent productivity effect (in percent) is then 0.0043/(0.01*0.0325).

This beneficial impact of information diffusion does not extend to other types of signals. Receiving signals from other (columns 2-4) or related (columns 5-6) markets at best has a mild effect on entry.[48] Matching export signals thus appear to be the main channel through which networks shape export behavior on the extensive margin of trade. To corroborate this result, we now turn to a series of robustness checks that relax several assumptions of our benchmark setting.

Endogenous export signals and linkages: To start with, we allow export signals $Y_{t-1}^{(jd)}$ and network linkages S_{t-1} to be correlated with outcome error ε_t. As explained in section 2.4, this form of endogeneity can arise due to common shocks that create a correlation in entry behavior among connected firms or induce firms to simultaneously seek out more linkages in domestic and foreign markets. To address a potential estimation bias from these channels, we employ two distinct strategies. First, we exploit the full network structure to identify firms that operate outside the network of seller i but are connected to immediate network peers j.[49] We then use lagged export decisions of these out-of-network firms as instruments for (potentially) endogenous entry decisions of immediate network peers j and estimate network effects via a two-stage least squares (2SLS) procedure. Second, we follow Arduini et al. (2015) and Qu et al. (2017) and estimate a dyadic network formation model.[50] Model estimates allow us to construct a selection correction term Ξ_{t-1} that accounts for any selection bias related to correlated formation decisions within domestic and foreign (export) networks.

Results of our instrumentation strategy and network selection correction are presented in appendix D.2.1 and D.2.2. Table 6 reports first- and second-stage results of our 2SLS estimation. As expected, export decisions of first- and second-order buyers appear to be strongly correlated (columns 1 and 3), underlining the relevance of the proposed instrument. Second-stage results (columns 2 and 4) indicate that matching signals remain a statistically significant determinant of entry, even after accounting for endogenous export signals. In fact, marginal effects of our IV

[48] Export starts within the European Economic Area (EEA) have been excluded from the sample as explained in section 3. Column 4 therefore represents a placebo test as signals from EEA countries shouldn't have any impact on export decisions in non-EEA markets. Columns 5 and 6 show the effect of signals from markets that share a border or have historic ties with export destination d. The mildly positive effect on entry points to a weak correlation in entry barriers across spatially proximate or culturally connected markets. The diffusion of export information within networks might therefore also contribute to spatially correlated entry behavior as observed by Albornoz et al. (2012) and Morales et al. (2019).

[49] As explained in section 2.4, we ensure that second-order buyers h and sellers i are not connected via higher-order linkages or located in close proximity to each other.

[50] A detailed overview of the formation model and the estimation procedure is presented in appendix D.2.2.

estimates exceed benchmark results by a factor of 1.5-2.5. Our instrument appears to correct for a downward bias, indicating that true network effects may be larger than suggested by our initial estimate. Adding Ξ_{t-1} as an additional regressor to account for selection bias from endogenous network linkages further corroborates this result as demonstrated in Tables 8 and 9. While selection correction terms are statistically significant, which implies that the underlying formation process is indeed endogenous, benchmark and IV estimates of network effects remain largely unchanged. The relevance of matching signals for the extensive margin of trade does not appear to be driven by endogeneity in the network formation process.

Model specification: Next, we compare benchmark estimates with alternative estimation approaches and fixed-effects specifications. In appendix D.2.3, we relax linearity assumptions imposed by our LPM-FE estimator and move to Logit and Probit models which allow marginal network effects to vary with characteristics of each seller. To avoid incidental parameter problems (IPP) that arise when combining these non-linear alternatives with high-dimensional fixed effects, we employ bias-adjusted non-linear logit and probit models developed by Fernández-Val and Weidner (2016) and Hinz et al. (2021) which mitigate IPP concerns and remain comparable to our linear estimates via average partial effects (APEs). Results presented in Table 10 showcase that APEs remain remarkably close to each other irrespective of which model is preferred.

The same is true for alternative fixed effects specifications presented in Table 11 in appendix D.2.4. Even in the most stringent specification (column 4) which features firm-year FE, destination-year FE, and firm-destination-FE, and absorbs any productivity differences across firms and firm networks, matching signals remain an important determinant of export participation.

Sample selection: In appendix D.2.5, we impose an alternative network cutoff value $\nu_{ij,t} = 5\%$ which drastically reduces the number of relevant network linkages. In this conservative setting, matching signals become an even rarer phenomenon but continue to be the dominant source of network effects as seen in Table 12. The marginal benefit of receiving a matching signal also becomes twice as large as our benchmark estimate.

Similar results are observed when restricting the sample to first-time exporters, whose export entry marks the first sale to a foreign market in the (observed) history of the firm. Table 13 in appendix D.2.6 shows that the marginal benefit of receiving a matching export signal exceeds our benchmark estimate by a factor of 2. Access to export experience via network linkages appears to be particularly important for first-time exporters.

Heterogeneous effects: The preceding results point to a large degree of heterogeneity in network effects among different firms and linkages. To understand the scope of this new mechanism for the extensive margin of trade, we need to explore how network effects are related to the characteristics of different firms, linkages, and geography. The results of this extensive exercise are presented in appendix D.2.7 and deliver two main messages: First, they demonstrate that the benefits provided by network linkages extend to a wide range of firms, characterizing network effects as a general pathway towards wider export participation. While individual firm and linkage characteristics can dampen or amplify the marginal effect of export signals, the overall scope of network effects does not appear to be limited to a select number of firms, linkages or sectors. Second, the results show that the diffusion of network benefits is not restricted to firms operating in close spatial proximity. This distinguishes our channel from classic agglomeration effects which largely accrue in the immediate vicinity of the firm.

4.2 Network channels

Table 4: Network effects and the margins of trade

	extensive margin		intensive margin			
	(1)	(2)	(3)	(4)	(5)	(6)
matching signals	0.0043***	0.0092*	-0.0006	0.0134	0.0117	0.1322
	(0.0011)	(0.0052)	(0.0084)	(0.0686)	(0.0201)	(0.0996)
IV		✓		✓		✓
selection correction		✓		✓		✓
R^2	0.093	0.093	0.415	0.412	0.531	0.843
Observations	469,770	453,532	94,318	90,445	143,107	226,877

Table 4 shows how export signals affect the entry probability and export volume of export starters and incumbents. Extensive margin regressions follow benchmark equation 7. Intensive margin regressions are based on equation 13. The full table is presented in appendix D.3. Column pairs 1-2 and 3-4 show the impact of signals on entry probability and export volume of export starters. Column pair 5-6 shows how export signals affect export volumes of incumbent exporters. All regressions include firm and destination-year FE and control for seller characteristics, demand, and productivity spillovers from network peers. Instruments and selection correction terms follow the procedure outlined in section 2.4. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

The intensive margin of trade: Our augmented entry framework allows network externalities to affect export participation by raising export profit or reducing the cost of entry. The positive effect of matching signals on entry documented in section 4.1 could therefore operate through either of the two channels. To link our empirical evidence to a distinct mechanism, it is useful to turn to the intensive margin of trade. If export signals indeed promote entry by raising export profits, both new and existing exporters should experience an increase in exports after receiving a matching signal.

To test this prediction, we follow Koenig et al. (2010) and estimate the impact of network effects on export volumes of starters and incumbents.[51] Details of the estimation procedure are presented in appendix D.3 and summarized here in Table 4. Columns 1 and 2 review results from the extensive margin of trade. As discussed, signals are associated with higher export participation, even after controlling for potential endogeneity of export signals and network linkages. This beneficial role of networks does not carry over to the intensive margin of trade. As seen in column pairs 3-4 and 5-6, export volumes of starters and incumbents do not respond to incoming matching signals. The same holds for other types of signals and productivity spillovers as seen in Tables 14 and 15. Matching signals therefore stimulate entry but appear to have no impact on firms after they have entered a foreign market. This pattern strongly suggests that domestic network effects operate through a sunk cost channel but leave variable costs and profits unchanged.

Demand uncertainty: While this interpretation is consistent with the empirical framework presented in section 2.2, we cannot rule out alternative explanations that lie outside of the current model. An important example is demand uncertainty. Our model assumes perfect foresight and thus abstracts from uncertainty as a determinant of entry. Uncertainty about demand in a foreign market can deter firms from entry, as firms do not know whether future profits are large enough to compensate for the sunk costs associated with market access. The observed benefit of matching signals for entry could therefore also be driven by a reduction in demand uncertainty, rather than a drop in entry-related sunk costs.

To disentangle these two competing interpretations, we compare the exit behavior of entrants with different signal exposure during the first three years after entry. If signals reduce demand uncertainty in market d, conditional on entry, signal-receiving firms are expected to have higher survival rates during the initial post-entry periods because they were more informed about demand when deciding to enter.[52] In appendix D.4, we test this idea empirically by tracking

[51] We define a firm as an incumbent if it has been exporting to a given market for two consecutive periods.

[52] Conditional on entry, signals are expected to increase survival rates because they inform firms about their true appeal in the foreign market. Firms that are discouraged from entry due to low demand will decide not to enter in the first place. Firms that enter without receiving any export signals on average are expected to be more likely to misjudge their appeal and thus more likely to exit in period $t + n$.

entrants during the first three periods after entry and ask whether matching signals received *prior* to entry have any impact on exit probabilities. Results are presented in Table 17. Across different time periods, model specifications and groups of entrants, we cannot find any evidence that suggests that matching signals lead to differential exit rates among entrants. While signals promote entry, they appear to be unrelated to subsequent firm behavior. This strongly suggests that signals operate through a sunk cost channel, rather than a reduction in demand uncertainty. To further corroborate this view, we investigate how the content of a signal affects the export decision of Belgian firms. To capture the content of signals empirically, we divide them into two groups. Firms that immediately exit a market after entry emit a 'bad' signal. Those that continue to serve the market beyond the initial entry period emit a 'good' signal. We thus interpret a firm's tenure after entry as a proxy for foreign market conditions. If signals affect entry by reducing uncertainty about foreign demand, bad signals should deter entry of connected firms, while positive signals should encourage it.[53] Empirical evidence presented in Figure 16e in appendix D.2.7 rejects this idea. Receiving an export signal, irrespective of whether it is good or bad, consistently leads to higher export participation.

The nature of network effects: What does this tell us about the nature of network effects? Leading frameworks in which signals reduce demand uncertainty among connected firms such as Fernandes and Tang (2014) show that signal effects are content-specific and simultaneously affect entry probabilities, export sales, and survival rates in foreign markets. In our setting, all matching signals facilitate market access but have no impact on post-entry sales or survival rates. Linking higher entry rates of signal-receiving firms to a reduction in demand uncertainty therefore lacks any empirical support. Instead, the available evidence clearly points to a sunk cost channel, consistent with the idea that signals promote entry by lowering the initial information cost associated with destination-specific import regulations and procedures. Network effects therefore represent a new determinant for the extensive margin trade that operates alongside other channels such as local agglomeration effects or learning about foreign demand.

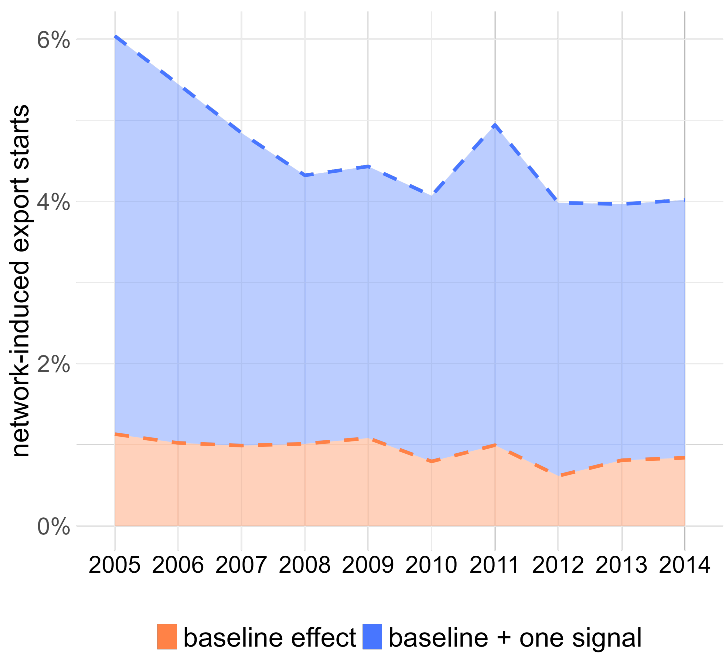
[53] Signals here are thought of as being part of a learning process in which firms update beliefs about foreign demand based on the export experience of their neighbors as in Fernandes and Tang (2014).

4.3 How important are network effects on aggregate?

Firms that receive export signals are more likely to become exporters. But how important is this channel on aggregate? To answer this question, we perform back-of-the-envelope calculations that allow us to compare network effects with other entry determinants and illustrate how many export starts can be directly attributed to the diffusion of export signals.

Network effects vs. productivity: We start by comparing marginal network effects to the most prominent entry determinant in the literature: firm productivity. As shown in columns 1 and 2 of Table 3, each additional matching signal raises the entry probability of a firm to a given destination by 0.4-0.9 percentage points. The effect of a single export signal is equivalent to a 13-30% increase in firm productivity. Access to export information thus has far-reaching consequences for the signal-receiving firm. While the individual effect of a matching signal is substantial, the low prevalence of signal diffusion means that the aggregate importance of network effects remains small. On average, firms receive a single matching signal every five years. The aggregate annual benefit of signal diffusion *on entry* is therefore equivalent to a uniform increase

Figure 4: Aggregate network effects – extensive margin



This figure depicts the aggregate impact of signal diffusion on the extensive margin of trade. It uses matching signal estimates reported in column 2 of Table 4. The y-axis indicates how many additional export starts at the firm-destination level are generated by matching signals compared to a setting with no signal diffusion. The orange line shows the number of entries under the current level of diffusion. The blue line presents estimates from a counterfactual in which each seller-destination pair receives one additional signal.

in productivity among all firms of 2.6-6.0%.

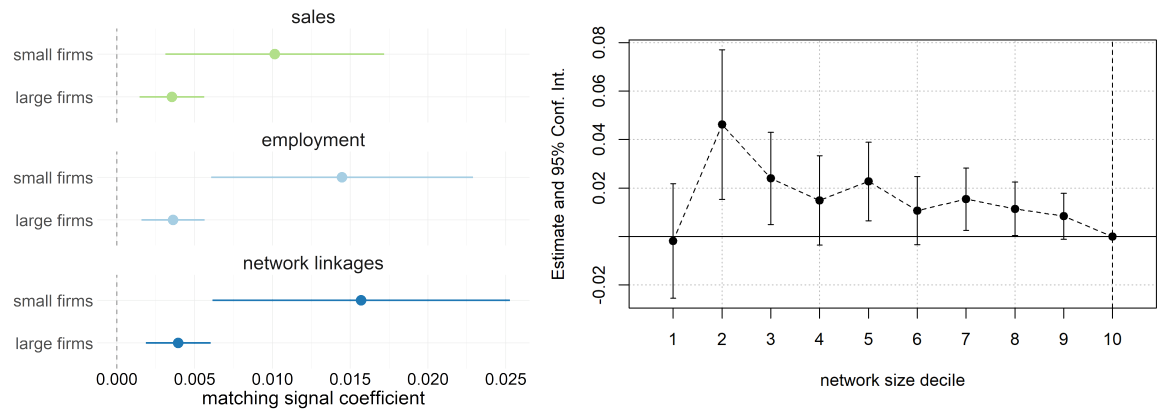
Network-induced entry: To quantify how many export starts are related to the diffusion of export signals, we compare the model-implied number of starts to a counterfactual where no signals travel along network linkages. Results of this exercise are presented in Figure 4. Eliminating the diffusion of signals reduces the annual number of export starts by 0.5 - 1%. This modest effect is largely driven by the low prevalence of signal diffusion. To illustrate this point, we construct a second counterfactual in which all firms receive an additional matching signal. Bridging the information gap by providing all firms with a minimum amount of information increases the total number of annual export starts by 3-5%. This result demonstrates that the potential benefits of network effects loom large, but are held back by a low prevalence of signal diffusion. Moreover, it opens the door for policy interventions to tap into the existing knowledge stock in domestic networks to achieve wider export participation – a point we discuss further in section 5.

4.4 Do network effects scale with network size?

Finally, we study how network effects contribute to the observed firm-size concentration on the extensive margin of trade. A large literature in international trade has documented that exporting is dominated by a small number of large and highly productive firms (Bernard et al., 2003; Mayer and Ottaviano, 2008). As seen in section 3.2, the very same firms also happen to have the largest networks (Figure 12) and receive the largest number of export signals (Figure 3a). This raises the question whether network effects further accelerate the concentration of export activity by providing large firms with a disproportionate amount of information, making it easier for them to expand into new markets.

Returns to network scale: To answer this question, we need to understand how marginal network effects are related to the size of the network they originate from. If the returns to receiving an export signal are constant or increasing in network size, export participation will mechanically be skewed towards firms with larger networks. Conversely, if the value of a signal decreases with network size, firms with smaller networks might benefit more from the diffusion of export information, despite receiving less information in absolute terms. Empirically, we test this idea by conducting two distinct exercises in which we compare how marginal network effects depend on the size of signal-receiving firms and their networks. Firm and network size in the Belgian data are strongly correlated (Bernard et al., 2022) which is why we use both as

Figure 5: Marginal network effects by firm and network size

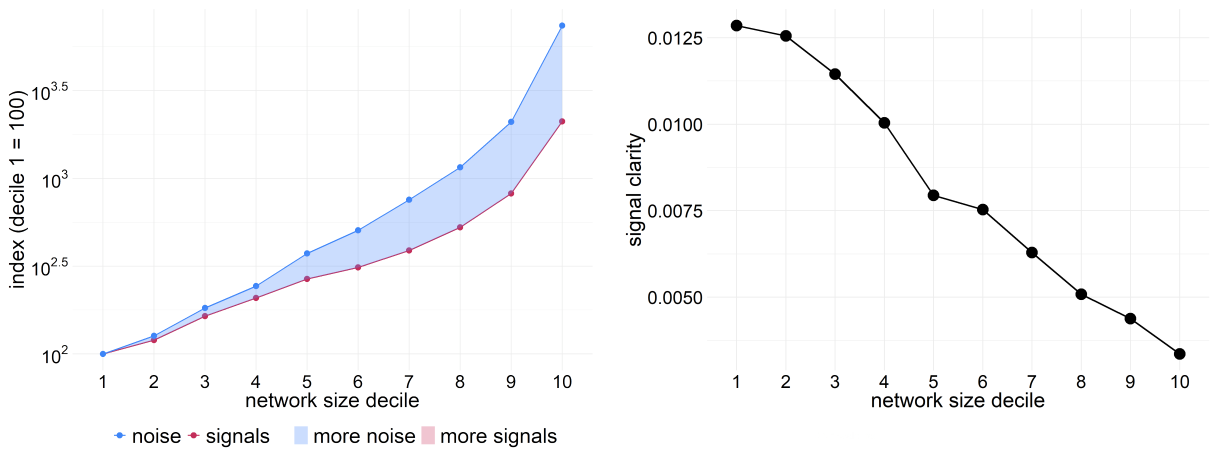


(a) subsample regressions (b) full sample with interaction term

proxies for network size. Results are summarized in Figure 5. In Figure 5a we divide sellers into two groups (small and large) based on their median sales, employment and the number of network linkages. We then run separate regressions for each subsample and plot matching signal coefficients. Irrespective of which size dimension is considered, we find that the marginal effects are consistently larger for small firms and networks. Signals emanating from small networks have a 2-4 times larger effect on entry compared to those originating in large networks. To further corroborate this result, we perform a second exercise in which we estimate separate network effects for each network size decile by introducing an interaction term to equation 7. Coefficients are plotted in Figure 5b and expressed relative to firms with the largest networks. Apart from firms with very small networks, marginal effects appear to decrease linearly in network size. This indicates that network effects exhibit decreasing returns to network scale.

What is causing decreasing returns? To rationalize this surprising result, two different explanations come to mind. First, the value of signals themselves might be subject to diminishing returns due to content overlap. Firms with access to large networks are more likely to receive the same signal multiple times and therefore respond less to each individual signal they receive. Unfortunately, our data does not allow us to observe the content overlap of multiple matching signals received by the same firm. While this prevents us from exploring this channel empirically, the low prevalence of signal diffusion means that most firms never receive more than one matching signal in the same period. Lower marginal effects in large networks are therefore unlikely to be the result of content overlap.

Figure 6: Network noise by firm size



(a) signal vs. noise growth rates (b) signal clarity

Instead, the marginal effect of signals might fall because of an increasing level of noise in large networks that makes it harder for firms to process incoming information. As explained in section 2.3, we define network noise as the share of linkages that do not yield any export signals. Figure 6a plots the average growth rates of export signals (red) and noise (blue) across different network size deciles. As expected, both lines are monotonically increasing given that larger networks naturally include a larger number of valuable and non-valuable linkages. The key insight, however, is that both linkage types grow at different rates. While growth rates are comparable across small networks, noise grows faster in medium-sized and large networks as indicated by the shaded blue area between both lines. Firms with large networks therefore receive more export signals on aggregate (signal intensity), but at the same time are exposed to disproportionate levels of network noise. As a result, the number of signals per linkage (signal clarity) decreases in larger networks as shown in Figure 6b.[54]

Signal clarity regressions: To test whether signal clarity can explain decreasing returns to network scale, we rewrite equation 7 and replace the number of export signals ($S_{t-1}Y_{t-1}^{jd}$) with the share of signals per linkage ($\bar{S}_{t-1}Y_{t-1}^{jd}$). Coefficient β now captures the impact of a unit

[54]These systematic differences in signal clarity across small and large networks are closely related to negative assortative matching in the underlying network formation process. Highly productive sellers on average interact with less productive buyers (Bernard et al., 2022). This inverse relationship between buyer and seller performance has direct implications for information diffusion. As shown in appendix B.5, when sellers become more productive they form more linkages, but the average share of exporters in their network falls as the network becomes large. The rising number of non-exporters do not contribute to the diffusion of export information and instead lead to increasing levels of network noise and lower signal clarity in large networks.

increase in signal clarity on the probability of entry. Since signal clarity is defined as the ratio of signals per linkage, β captures how much more effective an export signal becomes when it is received alongside less network noise. In other words, it describes how the marginal effect of a signal changes when reducing the size of the network it originates from.

The results of signal clarity regressions are presented in appendix D.5. In Figure 17, we plot signal clarity coefficients across network size deciles, to illustrate how firms with different networks respond to a reduction in network noise (increase in signal clarity). Two important results emerge. First, firms across size deciles benefit from lower levels of network noise and become more likely to enter foreign markets. This indicates that the efficacy of network effects in production networks depends on both the absolute amount (signal intensity) and signal-to-noise ratio (signal clarity) of transferred information. Second, firms with large networks respond more to a relative decrease in network noise than firms with small networks. We interpret this as direct evidence that noise is contributing to the decreasing returns in network scale observed above. Two identical firms that receive the same export signal do not benefit equally from the provided information if their networks differ in size, because they face different levels of network noise. Alternative explanations such as signal redundancy cannot generate this pattern, as they solely focus on content overlap and thus ignore any impact of network noise. We discuss the consequences of this finding in the next section.

5 Conclusion

In this paper, we show that who firms interact with in their domestic production network and what their peers know is crucial to understanding export behavior on the extensive margin of trade. This network perspective complements traditional notions of firm heterogeneity. While firm productivity remains a key driver of market entry, networks also play a crucial role by providing access to export-related information that lowers the cost of entry. The determinants of export participation are thus generated both within and outside of the firm. This insight has a number of important implications.

First, considering networks as determinants of entry helps us to rationalize firm behavior that is hard to reconcile with existing frameworks. An example is the existence of small exporters which often fail to meet theory-implied productivity thresholds that would justify their participation in foreign markets. Networks offer an alternative route to entry by allowing firms with experienced networks to enter despite insufficient levels of productivity.

Another example are entry decisions that defy a common rank order according to which firms should prioritize entry in markets that are large or easily accessible. In section 3.2, we document that during our sample period a larger number of Belgian firms chose to enter the Congolese market than the Chinese market – a clear violation of a theory-implied rank order. Even rich frameworks that use search and matching frictions to describe cross-border link formation (export entry) often implicitly assume a rank order of export markets. In these models, two identical firms with equally productive networks will typically choose to enter the same destinations, making it hard to rationalize entry decisions towards smaller markets, as seen in the Belgian context. Recognizing that production networks do not just coordinate the movements of goods and services but also expose firms to varying degrees of destination-specific information adds a new layer of granularity to the export decision process and helps us to better understand observed firm behavior.[55]

Second, network effects exhibit distinct features that clearly separate them from other entry determinants and force us to reconsider the role of firm size on the extensive margin of trade. Unlike productivity, which grows with firm size and raises export participation and profits across destinations, marginal network effects are destination-specific, do not extend beyond entry, and dissipate as firms and their networks become large. Decreasing marginal returns to network scale imply that small firms stand to gain more from each additional signal they receive. Our network channel therefore contributes to wider participation rather than increasing the existing firm-size concentration on the extensive margin of trade.

Observing decreasing returns also raises interesting questions about the optimal size of a network. If a network expansion adds noise to incoming export signals but benefits the firm in other dimensions, there could exist an optimal size threshold at which both effects are balanced. We are not convinced that learning about export markets is a first-order concern in the network formation process, which is why we disregard strategic network formation and treat network effects as a pure externality. Our work nonetheless highlights a potentially important indirect cost of unrestricted network expansions that goes beyond direct search and matching costs typically

[55] High entry rates to the Congolese market could for example point to a rich stock of market experience accumulated during the countries' colonial past that diffuses among connected firms and continues to promote entry today.

associated with the network formation process.

More generally, our results indicate that the scale effects in networks are closely related to the type of assortativity in the underlying formation process. Decreasing returns in our setting are linked to negative assortative matching in the production network. More productive sellers *on average* interact with less productive buyers. We show that this negative correlation between agent and peer characteristics creates noise and acts as an attenuating force for network effects. Conversely, if networks were subject to positive assortative matching, noise would decrease with network size, resulting in increasing returns to network scale. Knowledge of the direction of assortativity is therefore crucial when trying to understand the direction of scale effects in networks.

Finally, network effects offer policymakers a new lever to achieve broader export participation among domestic firms. In the face of an ever-evolving regulatory landscape, information frictions remain an important barrier to entry. To address this problem, export promotion agencies invest a considerable amount of time and resources to provide a *select group of members* with up-to-date market information, or organize costly matching events to connect domestic firms with foreign buyers. Our results suggest that domestic production networks can act as a powerful tool to provide similar benefits to *all domestic firms* in a potentially more cost-effective way. Instead of trying to directly link domestic firms to foreign buyers, policymakers could target the existing stock of export experience in domestic networks and provide platforms that facilitate a wider diffusion of available knowledge among domestic firms. We demonstrate that the potential benefits of this approach loom large but are currently held back by limited signal diffusion. Establishing more domestic linkages between experienced exporters and prospective entrants therefore presents a promising strategy, especially for small and medium-sized enterprises (SMEs) that stand to gain most from network externalities.

Taken together, our insights provide a new perspective on firm behavior on the extensive margin of trade but also raise important questions such as strategic considerations in formation of domestic and foreign linkages and the optimal size of domestic networks. These questions lie outside the scope of the current paper and require a more structural treatment of network linkages and firm outcomes. We believe our results provide important empirical evidence for this future avenue of research and will promote a stronger consideration of networks in international trade.

References

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

Ahn, J., A. K. Khandelwal, and S.-J. Wei (2011). The role of intermediaries in facilitating trade. *Journal of International Economics 84*(1), 73–85.

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

Alfaro-Ureña, A., I. Manelici, and J. P. Vasquez (2022). The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages. *The Quarterly Journal of Economics 137*(3), 1495–1552.

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

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

Arkolakis, C., F. Huneeus, and Y. Miyauchi (2023). Spatial production networks. Working Paper 30954, National Bureau of Economic Research.

Arkolakis, C., F. Huneeus, and Y. Miyauchi (2025). Production Network Formation, Trade, and Welfare. CESifo Working Paper 11833, CESifo GmbH, Munich.

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

Bernard, A. B., E. J. Blanchard, I. Van Beveren, and H. Vandenbussche (2018). Carry-along trade. *The Review of Economic Studies 86*(2), 526–563.

Bernard, A. B., E. Dhyne, G. Magerman, K. Manova, and A. Moxnes (2022). The origins of firm heterogeneity: A production network approach. *Journal of Political Economy 130*(7), 1765–1804.

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

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

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2010). Wholesalers and Retailers in US Trade. *American Economic Review 100*(2), 408–413.

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

Bernard, A. B. and Y. Zi (2022). Sparse production networks. Working Paper 30496, National Bureau of Economic Research.

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

Boucher, V., M. Rendall, P. Ushchev, and Y. Zenou (2024). Toward a general theory of peer effects. *Econometrica 92*(2), 543–565.

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

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

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (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. (2008). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review 98*(4), 1707–1721.

Chaney, T. (2014). The Network Structure of International Trade. *American Economic Review 104*(11), 3600–3634.

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

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

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

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

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

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

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

Duprez, C. and M. Nautet (2019). Economic flows between Regions in Belgium. *NBB Economic Review*, 1–19.

Eaton, J., M. Eslava, D. Jinkins, C. Krizan, and J. Tybout (2025). A search and learning model of export dynamics. *Journal of International Economics 157*, 104155.

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

Eaton, J., D. Jinkins, J. R. Tybout, and D. Xu (2022). Two-sided search in international markets. Working Paper 29684, National Bureau of Economic Research.

Eaton, J., S. S. Kortum, and F. Kramarz (2022). Firm-to-Firm Trade: Imports, Exports, and the Labor Market. Working Paper 29685, National Bureau of Economic Research, Cambridge, MA.

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

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

Fontaine, F., J. Martin, and I. Mejean (2023). Frictions and adjustments in firm-to-firm trade. CEPR Discussion Paper 18110.

Fujii, D., Y. Ono, and Y. U. Saito (2017). Indirect exports and wholesalers: Evidence from interfirm transaction network data. *Japan and the World Economy 44*, 35–47.

Ganapati, S. (2025). The modern wholesaler: Global sourcing, domestic distribution, and scale economies. *American Economic Journal: Microeconomics 17*(1), 1–40.

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

Goldsmith-Pinkham, P. and G. W. Imbens (2013). Social Networks and the Identification of Peer Effects. *Journal of Business & Economic Statistics 31*(3), 253–264.

Halleck Vega, S. and J. P. Elhorst (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., A. Stammann, and J. Wanner (2021). State Dependence and Unobserved Heterogeneity in the Extensive Margin of Trade. *mimeo*.

Hsieh, C.-S. and L. F. Lee (2016). A Social Interactions Model with Endogenous Friendship Formation and Selectivity. *Journal of Applied Econometrics 31*(2), 301–319.

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

Huang, H., K. Manova, O. Perello, and F. Pisch (2024). Firm Heterogeneity and Imperfect Competition in Global Production Networks. CESifo Working Paper 11305, CESifo GmbH, Munich.

Iyoha, E. (2023). Estimating Productivity in the Presence of Spillovers: Firm-level Evidence from the US Production Network. Working Paper 24-033, Harvard Business School, Cambridge, MA.

Johnson, R. C. and G. Noguera (2012). Accounting for intermediates: Production sharing and trade in value added. *Journal of International Economics 86*, 224–236.

King, G. and L. Zeng (2001). Logistic Regression in Rare Events Data. *Political Analysis 9*(2), 137–163.

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

Koenig, P., F. Mayneris, and S. Poncet (2010). Local export spillovers in France. *European Economic Review 54*(4), 622–641.

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

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

Liu, X., E. Patacchini, and Y. Zenou (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.

Martin, J., I. Mejean, and M. Parenti (2020). Relationship stickiness and economic uncertainty. *mimeo*.

Mayer, T. and G. I. P. Ottaviano (2008). The Happy Few: The Internationalisation of European Firms: New Facts based on Firm-level Evidence. *Intereconomics 43*(3), 135–148.

Mayer, T. and S. Zignago (2011). Notes on CEPII's Distances Measures: The GeoDist Database. Working paper 2011-25, CEPII.

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

Mion, G., L. D. Opromolla, and A. Sforza (2025). The Value of Managers' Export Experience: Lessons from the Angolan Civil War. *Review of Economics and Statistics 107*(2).

Morales, E., G. Sheu, and A. Zahler (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 E. L. Scott (1948). Consistent Estimates Based on Partially Consistent Observations. *Econometrica 16*(1), 1–32.

Panigrahi, P. (2021). Endogenous Spatial Production Networks Quantitative Implications for Trade and Productivity. CESifo Working Paper 9466, CESifo GmbH, Munich.

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

Patault, B. and C. Lenoir (2024). Customer Capital Spillovers: Evidence from Sales Managers in International Markets. *American Economic Journal: Applied Economics 16*(4).

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

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

A Endogenous network formation

To account for endogenous network formation, we introduce the network selection model of Arduini et al. (2015) and Qu et al. (2017) to our estimation procedure. This allows us to express network endogeneity as an unobserved shock to domestic production network S and export behavior $Y^{(id)}$ and correct for the selection bias resulting from correlated linkage formation and entry decisions. Formally, network formation is expressed by equation 8. Firms trade off the value of being linked to other firms and form linkages if

$$V(s_{ij,t} = 1) - V(s_{ij,t} = 0) > 0 = U_{ij,t} + \xi_{ij,t} \tag{8}$$

where $U_{ij,t}$ represents the linkage surplus and ξ is a random error term. The surplus is typically expressed as

$$U_{ij,t}(\theta) = \theta_0 + \boldsymbol{z_{i,t}\theta_1} + \boldsymbol{z_{j,t}\theta_2} + \boldsymbol{z_{ij,t}\theta_3} + \theta_4 A_{ij,t-1} \tag{9}$$

where coefficients $\boldsymbol{\theta_1}$, $\boldsymbol{\theta_2}$, $\boldsymbol{\theta_3}$ and θ_4 capture the impact of individual characteristics of firms i and j, characteristics of the dyad ij, and their past relationship status.[56] The inclusion of dyad characteristics $z_{ij,t}$ serves two important purposes. First, they control for key matching determinants such as bilateral distance or common language of buyers and sellers in Belgium. Second, they provide plausibly exogenous variation for our main outcome variable in equation 7. To give an example: In the Belgian context with 3 official languages, matching rates among domestic firms are expected to increase dramatically if both firms speak Flemish. Speaking Flemish, however, will not help these firms to enter non-European export markets like China. Dyadic controls like common language are therefore highly relevant for the formation of domestic linkages but at the same time satisfy exclusion restrictions for export decisions of the same firms. This feature allows us to construct a valid selection correction term that controls for unobserved correlation in the formation of domestic and foreign linkages as explained below.

If we assume that the random surplus component ξ is i.i.d. and follows a logistic distribution, we can write the linkage probability $s_{ij,t}$ as

$$P(s_{ij,t} = 1) = P(U_{ij,t}(\theta) + \xi_{ij,t} > 0) = \frac{e^{U_{ij,t}(\theta)}}{1 + e^{U_{ij,t}(\theta)}} \tag{10}$$

[56]Controlling for past relationship status is important because firm linkages in production networks tend to be very persistent (Martin et al., 2020) due to high search costs involved in the matching process.

To arrive at this expression, we assume that the conditional probability of i forming a link with j is independent from the decision to interact with another firm k. This implies that there is no strategic behavior in the network formation process that may characterize linkage formation in practice. Given the large network size considered in our setting[57] we believe this assumption is appropriate to render the problem computationally tractable. The resulting formation process still includes many important characteristics of production networks. Seller i can interact with multiple buyers j at the same time, demonstrate persistence in their linkage decision, and attach value to having business partners in close proximity.

While endogeneity in this context arises from unobserved shocks to network formation and export entry error terms ξ and ε, the timing of events in outcome equation 7 implies that we are mainly concerned about common shocks that have an immediate impact on domestic matching but only alter export decisions in the next period. Continuing the example from above, this would mean that the unexpected change in firm capacity is first deployed in the domestic market before being rolled out to prospective foreign destinations. Shocks affecting network formation and market access contemporaneously do not need to be considered as we only allow (endogenous) network effects to change firms' export decisions with a lag.

Before formalizing the correlation between formation and outcome errors, it is important to underline the dimensions at which both error terms operate. Network formation only considers firm-level characteristics of i and j whereas export decisions occur at the firm-destination level. This implies that our selection correction approach will only be able to capture correlated network and export behavior at the firm-level as formation errors $\xi_{ij,t-1}$ of seller i do not vary across seller export destinations d.

We start by collecting all network formation errors of seller i from dyadic regression 10 in row vector $\xi'_{i,t-1} = \{\xi_{ij,t-1}\}_{j \neq i}$. To relate formation errors to destination-specific outcome errors $\varepsilon_{id,t}$, each block of seller-specific error terms $\xi'_{i,t-1}$ is then repeated for each export destination seller i serves in year t. We denote the extended vector of formation errors as $\xi'_{i\{d\},t-1}$, where subscript d indicates that original formation errors have been repeated d-times for each seller i.[58]

[57] There are around 100,000 firms in the Belgian production network. Further steps to reduce the dimension of the formation process are discussed in appendix D.2.2.

[58] Assume there are two sellers A and B. Each form linkages with buyers 1 and 2 but serve a different number of export destinations. Seller A exports to China and India, seller B only exports to India. The destination-extended vector of formation errors for all sellers would thus be

The correlation in error terms can then be expressed as $(\varepsilon_{id,t}, \xi'_{i\{d\},t-1}) \sim i.i.d.(0, \Sigma_{\varepsilon\xi})$ where $\Sigma_{\varepsilon\xi} = \begin{pmatrix} \sigma^2_\varepsilon & \sigma'_{\varepsilon\xi} \\ \sigma_{\varepsilon\xi} & \Sigma_\xi \end{pmatrix}$, σ^2_ε is a scalar, $\sigma'_{\varepsilon\xi}$ and $\sigma_{\varepsilon\xi}$ are $(n_{ijd}-1)$-dimensional row and column covariance vectors and Σ_ξ is a $(n_{ijd}-1)$ by $(n_{ijd}-1)$ diagonal matrix with scalars σ^2_ξ on the diagonal. If we stack all row vectors of extended formation errors in a matrix we arrive at:

$$\Xi_{t-1} = \begin{bmatrix} \xi'_{1\{d\},t-1} \\ \xi'_{2\{d\},t-1} \\ \vdots \\ \xi'_{n\{d\},t-1} \end{bmatrix}$$

and we can decompose the outcome error as:

$$\varepsilon_t = \eta \Xi_{t-1} + \upsilon_t$$

where $\eta = \Sigma_\xi^{-1}\sigma_{\varepsilon\xi}$, $\sigma^2_\upsilon = \sigma^2_\xi - \sigma'_{\varepsilon\xi}\Sigma_\xi^{-1}\sigma_{\varepsilon\xi}$ and υ_t is independent of formation error ξ_{t-1}. Plugging the decomposed outcome error into equation 7 then yields

$$Pr\left(Y_t^{(id)} = 1\right) = Pr\left(\gamma' X_t^{(i)} + \delta' \bar{S}_{t-1} X_{t-1}^{(j)} + \beta S_{t-1} Y_{t-1}^{(jd)} + \psi_i + \psi_{d,t} + \eta \Xi_{t-1} > \varepsilon_t\right) \tag{11}$$

where $\eta\Xi_{t-1}$ describes the selection bias induced by endogenous network formation. If $\sigma_{\varepsilon\xi} \neq 0$, seller networks S_{t-1} become endogenous and network effects δ and β will be biased unless we control for Ξ_{t-1}.

To construct the selection correction term, we follow Arduini et al. (2015) and assume that outcome error ε_t is normally distributed. This allows us to use predicted linkage probabilities $p = P(s_{ij,t-1} = 1) = \frac{e^{U_{ij,t-1}(\theta)}}{1+e^{U_{ij,t-1}(\theta)}}$ from equation 10 and construct the selection correction term using a Heckman-type mills ratio:

$$\hat{\Xi}_{i,t-1} = \sum_{j \neq i} s_{ij,t-1} \frac{\phi(\Phi^{-1}(p))}{p} + (1 - s_{ij,t-1}) \frac{\phi(\Phi^{-1}(p))}{1-p} \tag{12}$$

where ϕ and Φ represent probability and cumulative density functions of a standard normal variable. The estimated selection correction term can then be used as an additional regressor in equation 11 to purge the outcome error of unwanted correlation from endogenous network for-

$\xi = (\underbrace{\xi_{A1}, \xi_{A2}}_{\text{China}}, \underbrace{\xi_{A1}, \xi_{A2}}_{\text{India}}, \underbrace{\xi_{B1}, \xi_{B2}}_{\text{India}})$.

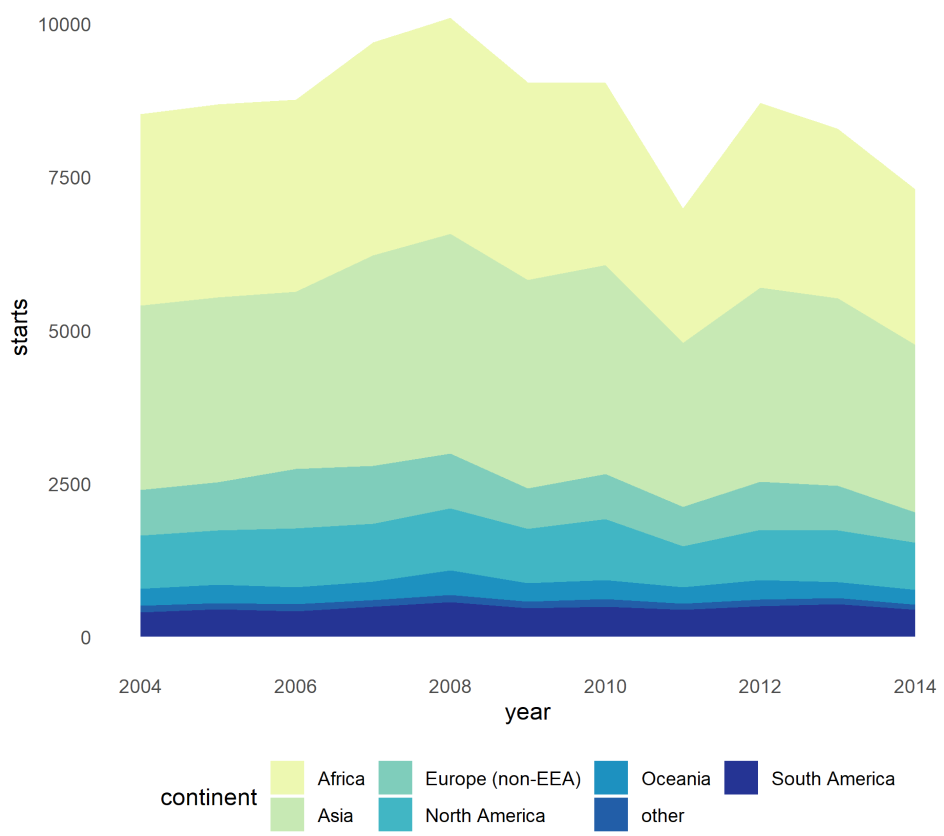
mation. A side effect of implementing this selection correction approach is the ability to directly test whether linkage endogeneity represents a concern when studying export entry decisions. If coefficient η is significantly different from zero, this would suggest that network formation is endogenous and controlling for selection is important to recover accurate estimates of network effects.

Results of the network formation procedure are reported in appendix D.2.2.

B Additional descriptives

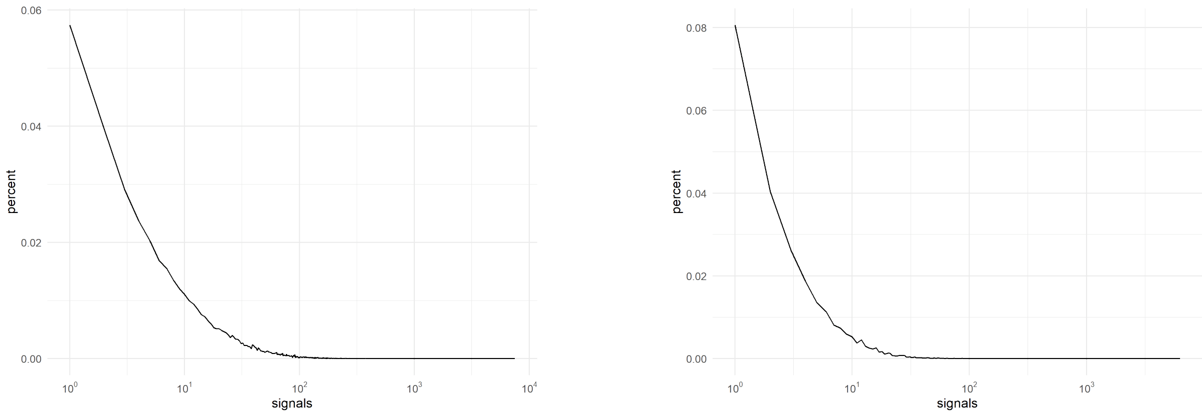
B.1 Export starts

Figure 7: Number of non-EEA starts by destination region



B.2 Export signals

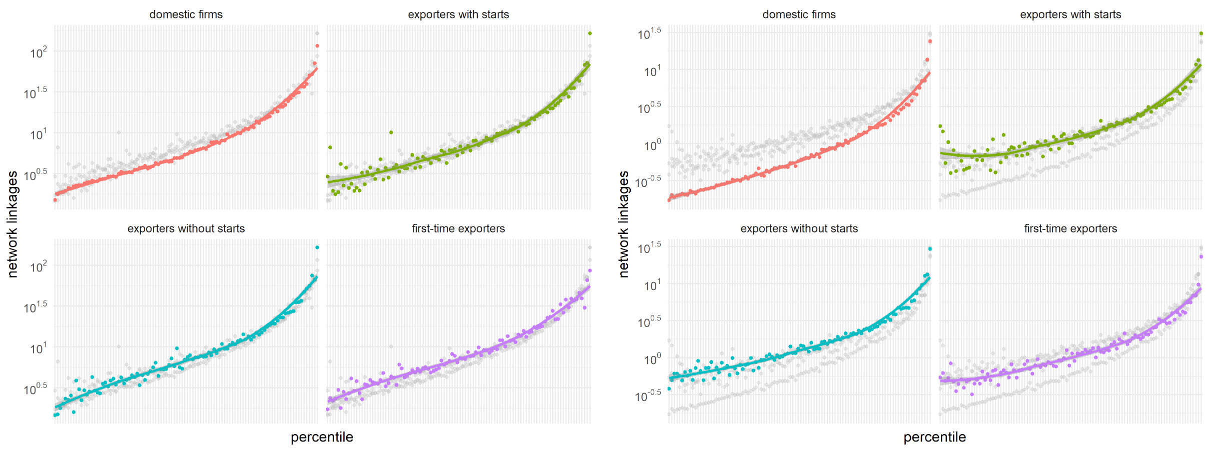
Figure 8: Distribution of firms receiving export signals



(a) any export signal (b) matching export signal

B.3 Network linkages

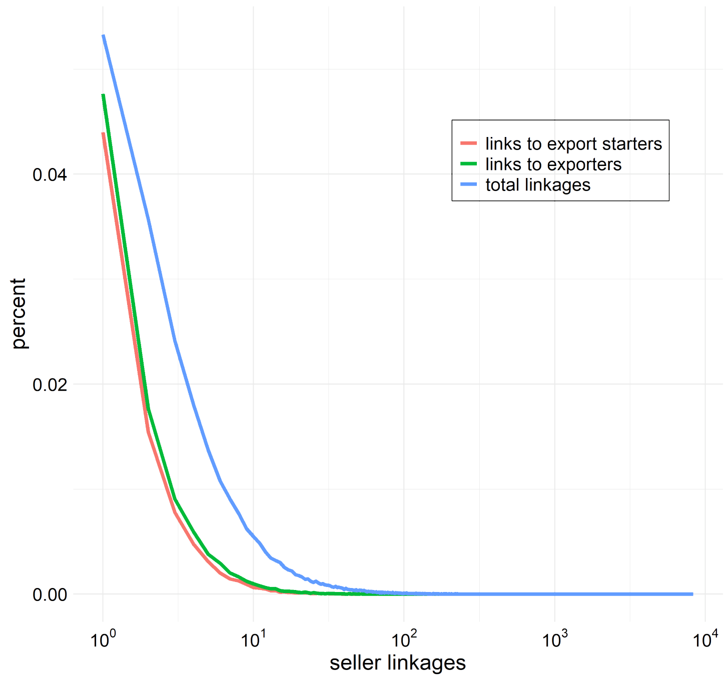
Figure 9: Seller linkages by TFP percentile



(a) all linkages (b) linkages to exporters

Note: This figure 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.

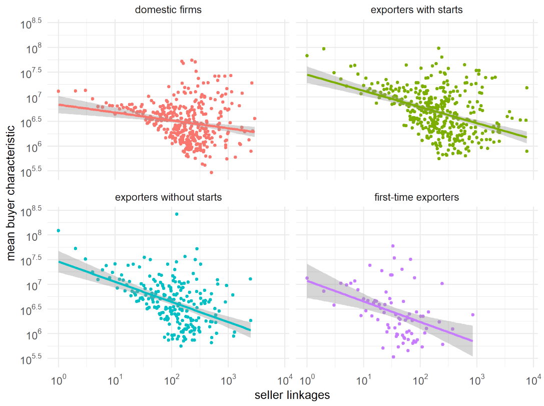
Figure 10: Distribution of seller linkages in 2014



Note: This figure shows the distribution of linkages to different types of buyers in 2014. Each line indicates what share of sellers interact with a certain number of buyers. The lines are based on a histogram and have been smoothed for visual purposes.

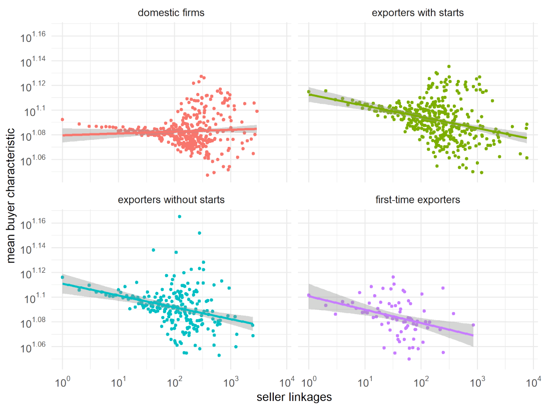
B.4 Network assortativity

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

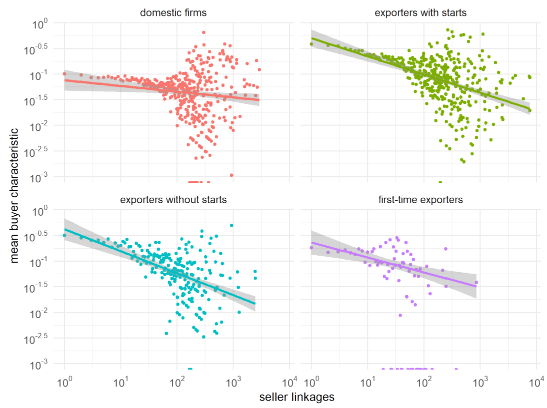


(a) average buyer sales

(b) average buyer employment



(c) average buyer productivity



(d) average buyer export start probability

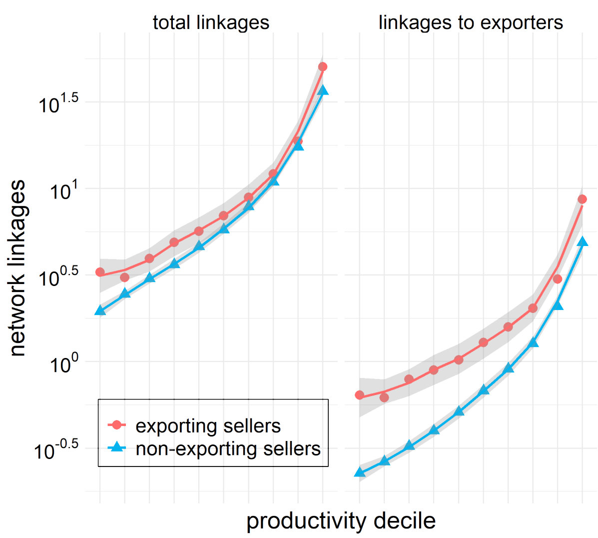
B.5 Stylized facts

In this section, we present two stylized facts to illustrate that network heterogeneity is not simply a primitive of firm productivity and that network benefits may not scale with network size. Both insights are important to understand how network externalities affect export participation decisions and to distinguish this channel from the existing literature.

Figure 12 illustrates our first stylized fact. It uses 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 decile. Both the left and the right side panel plot the number of network linkages of non-exporting (blue triangles) and exporting (red dots) sellers. The vertical distance between both lines represents the difference in network size. The panel on the left plots this difference for the entire buyer network of the seller while the right panel does the same for the subset of linkages which involve exporting buyers.

Focusing first on the left panel we observe two patterns. Firstly, across seller types the total

Figure 12: Buyer-Seller linkages by TFP decile



Note: This figure shows the average number of buyers that a seller in a given productivity decile interacts with. Seller productivity is computed using the approach of Levinsohn and Petrin (2003). Sellers are separated into non-exporters (triangles) and exporters (dots). The left panel shows linkages to any buyer, while the panel on the right, focuses on the subset of linkages involving buyers that export. The figure uses production network data of Belgian firms explained in detail in section 3.1.

number of linkages seems to increase in seller productivity. This pattern is common to production networks (Bernard and Zi, 2022). High productivity sellers are likely to attract more buyers because they can charge lower prices or offer better quality than their competitors. Secondly, we see that non-exporting and exporting sellers interact with a different number of buyers. Across productivity deciles, the average network size premium of exporting sellers (vertical distance between both lines) amounts to 27%. This indicates that exporting sellers overall seem to have more network interactions than domestic sellers, even after controlling for firm productivity.

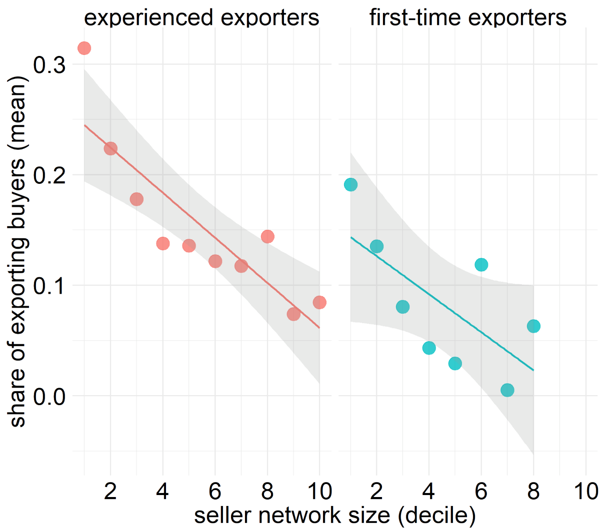
This difference increases dramatically when we move to the panel on the right, which focuses on seller linkages to exporting buyers. Network linkages with exporters appear to be much more common for firms that are exporters themselves. Domestic firms (non-exporting sellers) are much less likely to interact with exporters, which increases the average network size premium of exporting sellers to 103%. This significant wedge in network interactions of non-exporting and exporting sellers cannot be explained by seller productivity (which we condition on) or total network size (as seen on the left side panel), and holds for both incumbent and first-time exporting sellers (as shown in appendix figure 9). We summarize this finding as follows:

Stylized fact 1: *Comparing exporting to non-exporting sellers in the domestic network, we find that the average exporting seller has twice as many linkages to exporting buyers, even after controlling for seller productivity.*

This strong correlation between a seller's export status and linkages to exporters has two important implications. First, it indicates that network heterogeneity might be directly related to foreign market access. If interactions with exporters lead to productivity spillovers or a diffusion of export information, they could facilitate the export participation of connected firms. Second, it showcases that network heterogeneity is not a mere primitive of firm productivity. Firms with equal productivity interact with firms that have acquired different levels of export experience which may shape subsequent entry behavior.

A second stylized fact shows how networks change as they become larger. From Figure 12 we can see that more productive sellers have larger networks and are connected to a larger number of exporting buyers, as indicated by the upward sloping lines. This positive correlation between seller and network size, however, does not mean that the average performance of buyers in a large network is superior to buyers in a small network. In fact, seller productivity (and network size) is inversely related to average buyer productivity - a common pattern in production networks also known as negative assortative matching (Bernard and Zi, 2022; Bernard et al., 2022). Applied to our setting, where we focus on network linkages to exporters, it manifests as a negative correlation between a seller's network size and the share of exporting buyers in the network.

Figure 13: Seller network size and mean buyer export probability in 2014



Note: This figure shows the average share of exporting buyers for sellers in a given network size decile. The left panel shows linkages of sellers that have been exporting prior to 2014. The right panel focuses on the subset of sellers that started to export in 2014. The figure uses production network data of Belgian firms explained in detail in section 3.1.

Figure 13 shows this relationship for sellers with extensive export experience and those that just started to export for the very first time in 2014. In both cases, the share of exporting buyers decreases with network size. This pattern holds across years, seller types, and different buyer characteristics such as sales, employment, productivity, or the number of export starts, as shown in appendix B.4. This leads us to our second stylized fact:

Stylized fact 2: *Sellers with larger networks a larger number of linkages to exporters but the share of linkages to exporter falls with network size.*

This distinction between absolute and relative export exposure through network linkages is important because it suggests that network benefits might not scale with network size.

C Additional dataset information

C.1 Reporting thresholds for trade transactions

The reporting thresholds for trade flows in the Belgian data differ across intra-EU and extra-EU 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 and 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.

C.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.

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 and are 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) available at CEPII.[59] The latter includes information on bilateral relationships between 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, 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. *Export demand* captures changes in the demand for products that are responsible for the observed entry decisions of Belgian firms prior to the entry event. 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 *current* export value at HS6 product-level in *future* non-EEA export destinations of each firm i. This proxy captures the demand for each firm's products in market d prior to entry. We then aggregate this proxy 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.

[59]Specifically, we use version V202102 of the gravity dataset available at https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=8.

D Additional results

D.1 Benchmark regressions - full table

Table 5: Benchmark results - signal type

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0043***					
	(0.0011)					
non-matching signals		0.0000				
		(0.0001)				
total signals			0.0000			
			(0.0001)			
EEA signals				0.0000		
				(0.0001)		
border signals					0.0003	
					(0.0004)	
history signals						0.0002***
						(0.0001)
log employment	0.0476***	0.0477***	0.0477***	0.0477***	0.0477***	0.0476***
	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)
log wage	0.0127**	0.0127**	0.0127**	0.0127**	0.0127**	0.0127**
	(0.0060)	(0.0060)	(0.0060)	(0.0060)	(0.0060)	(0.0060)
log TFP	0.0325***	0.0325***	0.0325***	0.0325***	0.0325***	0.0326***
	(0.0040)	(0.0040)	(0.0040)	(0.0040)	(0.0040)	(0.0040)
log export demand	0.0179***	0.0179***	0.0179***	0.0179***	0.0179***	0.0179***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
border dummy	0.0694***	0.0694***	0.0694***	0.0694***	0.0694***	0.0694***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
history dummy	-0.0140***	-0.0140***	-0.0140***	-0.0140***	-0.0140***	-0.0140***
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0024)
export propensity	0.1702***	0.1701***	0.1701***	0.1701***	0.1701***	0.1701***
	(0.0115)	(0.0115)	(0.0115)	(0.0115)	(0.0115)	(0.0115)
import dummy	0.0980***	0.0982***	0.0982***	0.0982***	0.0982***	0.0982***
	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)	(0.0034)
log peer size	-0.0005	-0.0004	-0.0004	-0.0003	-0.0003	-0.0004
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
peer TFP	0.0004	0.0003	0.0004	0.0003	0.0003	0.0004
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0027)	(0.0027)
firm FE	✓	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓	✓
R^2	0.0926	0.0926	0.0926	0.0926	0.0926	0.0926
Observations	469,770	469,770	469,770	469,770	469,770	469,770

This table shows regression results of estimating equation 7 with a LPM-FE. Each column shows the marginal effect of receiving a different type of export signal on a seller's probability to start exporting. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2 Benchmark robustness

To assess the validity of our benchmark estimates presented in section 4.1, we perform five sets of robustness checks. First, sections D.2.1 and D.2.2 address endogeneity concerns related to endogenous export signals and network linkages. Second, sections D.2.3 and D.2.4 test alternative model specifications and estimation approaches. Third, sections D.2.5 and D.2.6 study how sample selection affects our network estimates. Next, section D.2.7 explores how network effects are shaped by heterogeneity in linkage and peer characteristics. Finally, section D.2.7 studies how geography shapes the diffusion of network effects.

D.2.1 Robustness: Endogenous export signals

Table 6: Endogenous export signals - 2SLS

Dependent variable: IV stages:	matching signals First	export starts Second	matching signals First	export starts Second
Variables				
second-order signals	0.2800***		0.3836***	
	(0.0606)		(0.0764)	
matching signal		0.0092*		0.0064**
		(0.0052)		(0.0028)
Peer characteristics	✓	✓		
Destination experience	✓	✓	✓	✓
Firm characteristics	✓	✓		
firm FE	✓	✓		
destination-year FE	✓	✓	✓	✓
firm-year FE			✓	✓
R^2	0.573	0.093	0.663	0.248
Observations	469,619	469,619	828,710	828,710

This table shows results of a 2SLS regression of equation 7. Endogenous export signals are instrumented by second-order signals. Columns 1 and 3 show first-stage results, columns 2 and 4 second-stage results. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2.2 Robustness: Endogenous network linkages

This section provides additional details on how to obtain selection correction terms from the Belgian network data. A detailed derivation of these selection terms is presented in appendix A. Here we focus on the estimation procedure. Readers primarily interested in how the selection correction affects benchmark estimates of network effects can proceed directly to tables 8 and 9.

How to construct selection correction terms: To begin with, we need to estimate the dyadic formation model outlined in equation 9, which poses two distinct challenges. A first challenge relates to the large sample size n. As the model requires us to estimate the linkage probability between any pair of firms operating in Belgium, each firm in theory considers all $n-1$ other firms as candidates for establishing a linkage. Including all $n \times (n-1)$ firm pairs in our dyadic formation model is not only computationally infeasible, since our sample includes around 100,000 firms per year, but also highly unrealistic, as firms are unlikely to consider the entire population of firms as matching candidates when searching for an individual business partner. A second challenge is that even if the size of candidate sets becomes computationally feasible, the exact candidate set a firm considered during the matching process remains unobserved. Both challenges require additional assumptions which we discuss in turn.

To reduce the dimension of the problem, we impose several restrictions on the $n \times n$ firm-to-firm interaction matrix. Instead of treating all $n-1$ firms as potential matching candidates for an observed linkage $s_{ij,t} = 1$, we only consider firms as candidates if they operate in the same 4-digit NACE industry as the actual match, and have interacted with firms in the same 4-digit NACE sector as firm i. All candidates that do not meet these criteria are dropped from the candidate set as if they were never considered as potential business partners for the observed firm-to-firm linkage. This two-sided sector-specific restriction creates a distinct candidate set for each observed linkage and significantly reduces the size of potential candidates. On average, there are 60 candidates per observed match.

While this selection potentially introduces some error by ruling out candidates that firm i actually considered during the matching process, we believe that restricting candidate sets to the sector of the actual match j is intuitive and is expected to preserve the majority of true candidates.

Since we do not know which of the 60 candidates firm i considered during the matching process, we take a random sample of $n^{random} = \{1, 5, 10, 20\}$ candidates and estimate the dyadic formation model with a logit model for a given draw of n^{random} candidates plus the actual match.

Results of the dyadic formation model for a given set size are presented in table 7. The model controls for firm size, productivity, and the import status of each individual firm i and j, and adds a set of dyadic controls that capture whether the two firms speak the same language, their bilateral distance, and whether they had a linkage in the previous period.

Our primary goal is to use the estimated formation coefficients to predict linkage probability $p = P(S_{ij,t-1} = 1)$ and compute selection correction terms $\hat{\Xi}_{i,t-1}$ as shown in equation 12 in appendix A. We therefore refrain from interpreting them as causal. To mitigate concerns that

Table 7: Network formation – linkage probability

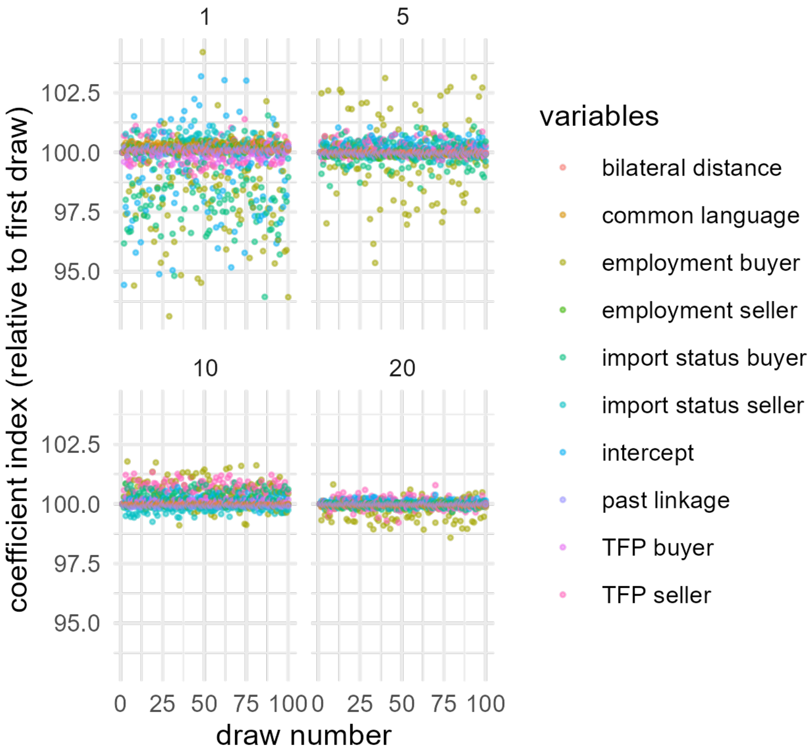
Candidates per match	n=1	n=5	n=10	n=20
Constant	0.229***	-0.654***	-1.03***	-1.47***
	(0.012)	(0.009)	(0.009)	(0.009)
log empl_i	0.158***	0.159***	0.172***	0.199***
	(0.0005)	(0.0004)	(0.0004)	(0.0004)
log empl_j	-0.010***	0.007***	0.015***	0.018***
	(0.0007)	(0.0006)	(0.0005)	(0.0005)
log TFP_i	0.058***	0.051***	0.042***	0.026***
	(0.0006)	(0.0005)	(0.0005)	(0.0005)
log TFP_j	-0.111***	-0.143***	-0.153***	-0.149***
	(0.0008)	(0.0007)	(0.0006)	(0.0006)
importer_i	0.240***	0.170***	0.181***	0.210***
	(0.002)	(0.001)	(0.001)	(0.001)
importer_j	0.040***	0.060***	0.101***	0.161***
	(0.002)	(0.001)	(0.001)	(0.001)
same language	0.601***	0.576***	0.560***	0.546***
	(0.002)	(0.001)	(0.001)	(0.001)
bilateral distance	-0.021***	-0.023***	-0.023***	-0.023***
	(2.04×10^{-5})	(1.79×10^{-5})	(1.73×10^{-5})	(1.69×10^{-5})
past link	6.11***	6.23***	6.27***	6.32***
	(0.008)	(0.004)	(0.003)	(0.003)
Pseudo R^2	0.391	0.424	0.429	0.432
Observations	14,421,866	38,792,114	64,154,714	104,598,865

This table shows how firm characteristics affect the linkage formation in the Belgian production network. Results are based on equation 8 and estimated via a logit model. Each column shows results for a sample in which each firm selects one peer from a set of n randomly drawn candidates.

selection correction terms might be driven by the chosen candidate set size n^{random}, we employ the rare events correction of King and Zeng (2001) which accounts for the different ratios of events ($s_{ij,t} = 1$) to non-events ($s_{ij,t} = 0$) across set sizes. As shown in table 7, coefficient estimates remain stable across sets of different sizes. Moreover, figure 14 illustrates that variation within a given set n^{random} due to the random sampling of candidates remains minimal. Across draws within the same set size, selection correction terms $\hat{\Xi}_{i,t-1}$ are therefore highly correlated.

Benchmark estimates with selection correction: To control for unobserved shocks that affect both the formation of domestic and foreign linkages, we add selection correction term $\hat{\Xi}_{i,t-1}$ as an additional regressor to benchmark equation 7. Results are presented in table 8. Despite the fact that selection correction terms are statistically significant, indicating that the underlying formation process is indeed endogenous, matching signals remain close to our benchmark estimates. A selection bias from endogenous network linkages does not appear to be a major concern in our setting.

Figure 14: Variation of network formation coefficients by candidate set size



Note: The figure plots the results of network formation equation 9. For each candidate set size n^{random}, we draw 100 random samples and display the resulting coefficients relative to the coefficient value of the first draw, which we index to 100.

Table 8: Endogenous network linkages - selection correction

Candidates per match:	baseline	n=1	n=5	n=10	n=20
Variables					
matching signals	0.0042***	0.0041***	0.0041***	0.0041***	0.0041***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
selection correction		0.0060***	0.0062***	0.0061***	0.0059***
		(0.0013)	(0.0014)	(0.0014)	(0.0015)
Peer characteristics	✓	✓	✓	✓	✓
Firm destination experience	✓	✓	✓	✓	✓
Firm characteristics	✓	✓	✓	✓	✓
firm FE	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓
R^2	0.093	0.093	0.093	0.093	0.093
Observations	445,620	445,620	445,620	445,620	445,620

This table shows results of estimating equation 11 which accounts for endogenous network formation via a selection correction term. The selection correction term is based on the dyadic network formation model outlined in equation 9 and calculated for several buyer candidate sets which differ in size. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

Table 9: Extensive margin robustness – endogenous signals and linkages

Source of endogeneity	baseline	signals	linkages	both
matching signals	0.0043***	0.0092*	0.0041***	0.0092*
	(0.0011)	(0.0052)	(0.0011)	(0.0052)
IV		✓		✓
selection correction			✓	✓
firm FE	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓
R^2	0.093	0.093	0.093	0.093
Observations	469,770	469,770	453,532	453,532

This table shows how export signals affect the entry probability of Belgian firms. Results are based on benchmark equation 7. All regressions control for seller characteristics, seller destination experience, and externalities based on network characteristics. Instruments and selection correction terms follow the procedure outlined in section 2.4. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2.3 Robustness: Nonlinear models

Table 10: Robustness - Nonlinear models

Model	LPM-FE	Logit-FE	Logit-FE-IPP	Probit-FE	Probit-FE-IPP
Coefficient for *matching signal*	0.0043***	0.0348***	0.0346***	0.0187***	0.0186***
	(0.0011)	(0.0059)	(0.0060)	(0.0034)	(0.0034)
APE for *matching signal*	0.0043***	0.0050***	0.0052***	0.0049***	0.0049***
	(0.0011)	(0.00089)	(0.00090)	(0.00091)	(0.00091)
firm FE	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓
Observations	469,770	475,915	475,915	475,915	475,915

This table compares regression results of equation 7 for different linear and non-linear models. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2.4 Robustness: Alternative fixed effect specifications

Table 11: Robustness - Fixed effects

Model:	(1)	(2)	(3)	(4)
Variables				
matching signals	0.0043***	0.0052***	0.0043***	0.0041***
	(0.0011)	(0.0008)	(0.0010)	(0.0010)
Peer characteristics	✓			
Firm destination experience	✓			✓
Firm characteristics	✓			
firm FE	✓			
destination-year FE	✓	✓	✓	✓
firm-year FE		✓	✓	✓
firm-destination FE			✓	✓
R^2	0.093	0.235	0.333	0.343
Observations	469,770	852,955	836,506	811,878

This table compares regression results of equation 7 for different fixed effect specifications. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2.5 Robustness: Alternative network threshold

Table 12: Robustness - 5% network threshold

Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
matching signals	0.0088***					
	(0.0022)					
non-matching signals		0.0001				
		(0.0002)				
total signals			0.0002			
			(0.0002)			
EEA signals				0.0001		
				(0.0005)		
border signals					0.0015*	
					(0.0009)	
history signals						0.0004*
						(0.0002)
Peer characteristics	✓	✓	✓	✓	✓	✓
Firm destination experience	✓	✓	✓	✓	✓	✓
Firm characteristics	✓	✓	✓	✓	✓	✓
firm FE	✓	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓	✓
R^2	0.094	0.094	0.094	0.094	0.094	0.094
Observations	360,253	360,253	360,253	360,253	360,253	360,253

This table compares regression results of equation 7 using a 5% buyer sourcing threshold to define relevant network linkages. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2.6 Robustness: First-time exporters

Table 13: Robustness - First-time exporters

No export activity before:	baseline	2003	2006	2012
Variables				
matching signals	0.0043***	0.0092***	0.0079***	0.0112
	(0.0011)	(0.0028)	(0.0031)	(0.0077)
Peer characteristics	✓	✓	✓	✓
Firm destination experience	✓	✓	✓	✓
Firm characteristics	✓	✓	✓	✓
firm FE	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓
R^2	0.093	0.155	0.153	0.299
Observations	469,770	136,017	109,018	23,611

This table compares regression results of equation 7 focusing on sellers with no export experience at the firm-level up to the indicated year. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.2.7 Robustness: Signal heterogeneity

In this section, we explore how network effects are related to the characteristics of different firms, linkages, and geography. Results are presented in different panels in figure 16. All panels depict matching signal coefficients obtained from equation 7 and are estimated via LPM-FE. The only difference to our benchmark specification is that matching signal are disaggregated by linkage and peer characteristics to assess the impact of network heterogeneity. The disaggregation exercise in each panel uses the following definitions:

- Panel (a) uses three approaches to separate linkages into strongly and weakly dependent. First, by ranking all buyers j based on their sourcing share from seller i. Buyers above the median rank are then defined as more dependent and vice versa. As an alternative, we use the observed sourcing shares to define strongly dependent buyers as those that source at least 50% (90%) of domestic inputs from seller i.

- Panel (b) defines linkage persistence as the number of consecutive years a seller-buyer pair ij have interacted with each other. Incoming export signals are then assigned according to the maturity of linkage ij in year t. To ensure all linkage maturities can be observed in our sample, regressions only consider entry decisions after 2006.

- Panel (c) considers export signals received from backward, forward, and mixed linkages. The three linkage types reflect the relationship of buyers and sellers in the production network. Backward linkages capture signals received from buyers, forward linkages capture signals received from suppliers, and mixed linkages capture signals received from firms that simultaneously act as buyers and sellers for firm i. The third regression in this panel excludes export starts of wholesalers by dropping firms operating in NACE sectors 45, 46, and 47 from the sample.

- Panel (d) disaggregates incoming export signals by peer size. We define large and small firms based on sales and use median sales as the cutoff value.

- Panel (e) studies the credibility of export signals in three distinct ways. First, by investigating whether peer entries are persistent or they immediately leave the market in $t + 1$. Second, by checking whether exports to the new destination account for more than 1% of total peer exports in that year. Third, by examining whether exports to the new market account for more than 1% of total exports in the firm's 4-digit NACE sector.

- Panel (f) finally uses NACE sectors 45, 46, and 47 to identify wholesalers and separately counts export signals originating from wholesaler and non-wholesaler networks.

We start by comparing how the strength of B2B linkages shapes the impact of export signals. To this end, panel (a) plots the estimated coefficients of matching signals that have been received via different linkage types. While the seller-specific rank of individual buyers does not seem to matter for the strength of network effects, receiving a signal from buyers which rely on a single seller for a majority of their sourcing (B2B sourcing shares above 50%) appears to have a stronger impact on subsequent entry decisions of sellers. While signals received from buyers with more diversified sourcing strategies still facilitate entry, this result suggests that linkage dependency amplifies the impact of export signals.

In panel (b) we study how the duration of the B2B relationship affects network effects. Although network linkages are often sticky, given the non-negligible fixed costs involved in identifying suitable business partners (Martin et al., 2020), we find that both new and persistent linkages facilitate export entry. This suggests that sellers not only respond to signals received from trusted sources but also remain open to insights from new business partners.

A last linkage characteristic which we investigate in panel (c) is the direction of signal diffusion. While our analysis focuses on backward linkages meaning sellers receive signals from their buyers, network effects might also arise from forward (the network of suppliers) or mixed linkages where firm i simultaneously acts as a buyer and a supplier for peers j. Coefficients obtained from benchmark equation 7 with weak and strict FE suggest that all three linkage directions can facilitate entry. However, once we remove export signals emitted by wholesalers (NACE sectors 45, 46 and 47) from the sample, only signals diffusing along backward linkages continue to promote entry. As discussed in section 3.1, wholesalers provide highly specialized services to their clients such as warehousing, transport and logistics. These services are often destination-specific and therefore likely contribute to a diffusion of export information along forward or mixed linkages. This direction of diffusion in which buyers learn from sellers, however, does not hold of other types of firms. The evidence presented here supports this view. Signals from non-wholesalers have no impact on entry, if they diffuse along forward or mixed linkages. While these alternative modes of diffusion can promote export participation for a specific subset of network linkages, they are not as prevalent as the far-reaching benefits provided by signal diffusion along backward linkages.

In panels (d), (e), and (f) we study the impact of peer heterogeneity on network effects. Panel (d) starts by investigating the role of peer size. Our findings show that export signals originating from small and large firms both matter for the observed conducive impact on export entry, but also reveal a substantial degree of homophily in the underlying network effects. While foreign market access of large firms is disproportionately driven by interactions involving other large firms, the opposite is true for small firms. Our empirical framework is not equipped to uncover the underlying mechanism at play, but suggests that network effects are linked to the degree of homophily within buyer-seller pairs.

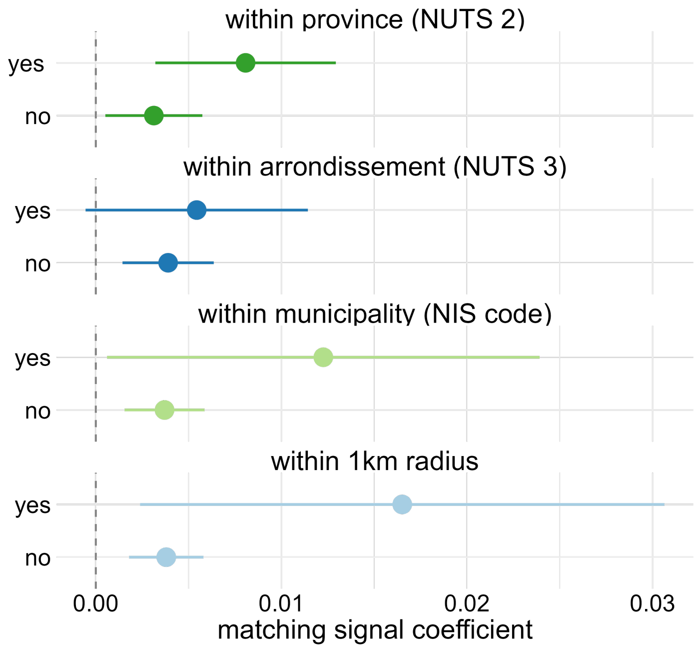
Next, we investigate whether the credibility of export signals shapes the entry behavior of sellers. Results are presented in panel (e). Our first disaggregation accounts for the export behavior of network peers one year after they emitted an export signal. Cases in which peers immediately leave the foreign market after their initial entry could signal unfavorable market conditions and nullify the positive impact of the emitted signal for the receiver. Our results do not corroborate this claim. Post-entry export behavior or peers appears to have no impact on the strength network effects. Conversely, we do find evidence that credible signals, proxied by the share of export volume in total peer or sectoral exports, does seem to have a larger impact on the receiver. While our benchmark analysis treats each incoming signal as a uniform piece of information, firms do seem to differentiate between different incoming signals.

In panel (f), we return to the role of wholesalers. Trade intermediaries have been shown to facilitate foreign market access by initially allowing domestic firms to bypass entry barriers by exporting through wholesalers before entering markets directly (Connell et al., 2019). Our framework generalizes this idea by considering all network interactions, including those with non-wholesalers, as a source of diffusion of valuable export information that promotes export entry. To showcase the general nature of this channel, we distinguish between wholesaler and non-wholesaler linkages and investigate how export signals from each group affect export entry of connected sellers. Our results reveal that non-wholesalers signals contribute consistently to the expansion of export activity on the extensive margin of trade, whereas wholesaler signals predominantly facilitate export entry of other wholesalers. The diffusion of export promoting information is therefore driven by a wide range of firms, which stresses the importance of considering the entire production network when estimating network effects.

Finally, we study the spatial range of signal diffusion to rule out that estimated network effects merely reflect agglomeration economies from buyers in close geographic proximity. Results of this final robustness check are presented in Figure 15 in appendix D.2.7. Reassuringly, we find that signals of buyers located outside the seller's province are equally conducive to export entry as signals originating from buyers located within the same province. This result has two important implications. First, it showcases that our network effects capture a general diffusion mechanism that is not limited to the geographic confines of a province, city, or street as in the preceding spillover literature (Koenig et al., 2010; Fernandes and Tang, 2014; Bisztray et al., 2018). Second, it rules out that estimated network effects are merely the result of labor movements between firms.[60] Evidence from Belgian commuter flow surveys shows a strong preference to reside in close proximity to the workplace as 85% of commuters do not cross a provincial border to go to work (Duprez and Nautet, 2019). Network effects extending beyond provincial borders as shown in Figure 15 are therefore unlikely to be driven by labor movements between buyers and sellers, as the majority of people do not seek employment outside their province.

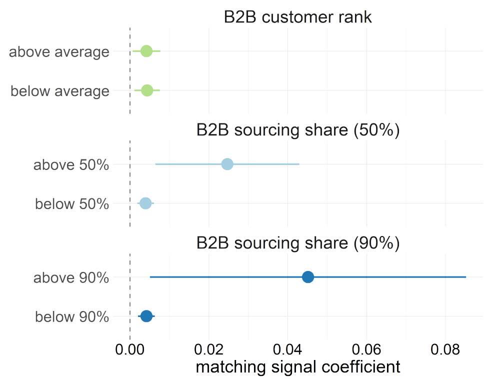
[60] For recent work studying this channel see Choquette and Meinen (2015) and Patault and Lenoir (2021).

Figure 15: Geography of matching signals

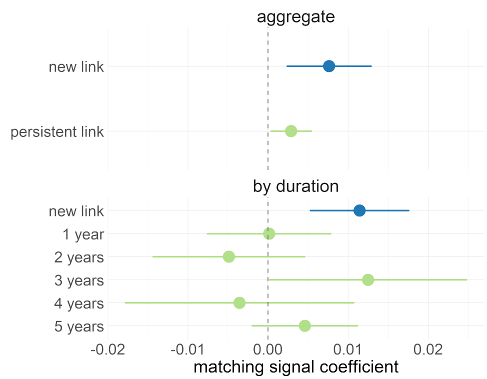


Note: This figure plots estimation results from a version of equation 7 in which export signals are disaggregated by the geographic location of the buyers and sellers. Belgium contains 11 provinces (NUTS 2 code), 44 arrondissements (NUTS 3 code), and around 590 municipalities (5-digit NIS code).

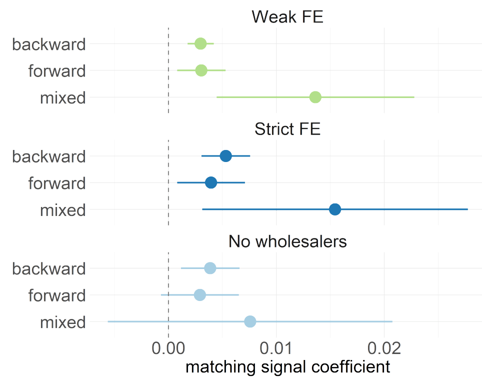
Figure 16: Matching signal heterogeneity



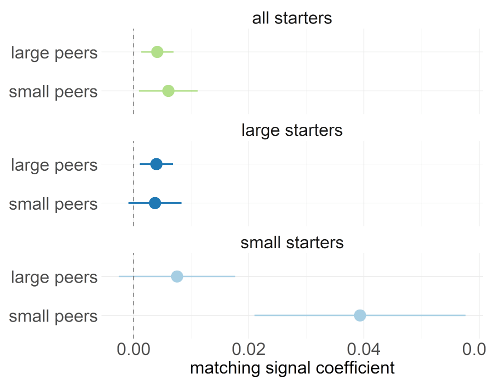
(a) linkage dependency



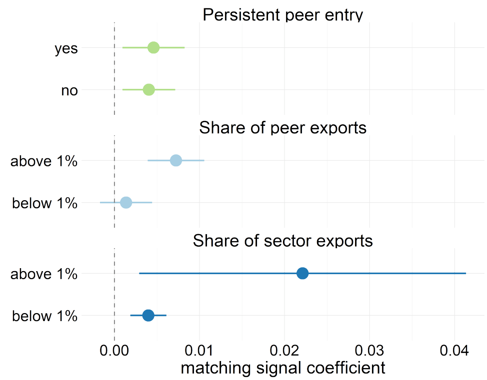
(b) linkage persistence



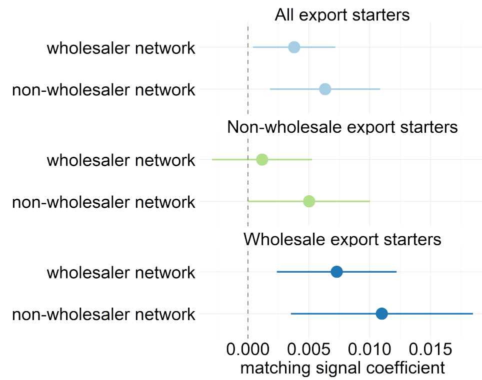
(c) linkage direction



(d) peer size



(e) peer credibility



(f) peer sector

D.3 Intensive margin

In this section, we present results that show how network externalities affect the intensive margin of trade. We follow Koenig et al. (2010) and estimate the following log-linearized equation via OLS

$$\exp_{id,t} = \boldsymbol{a}'\boldsymbol{x_{id,t}} + \rho D_{id,t-1} + \boldsymbol{\delta}'\sum_j \bar{s}_{ij,t-1}\boldsymbol{x_{j,t-1}} + \beta\sum_j s_{ij,t-1}y_{jd,t-1} + FE_i + FE_{d,t} + \varepsilon_{id,t} \quad (13)$$

where $\exp_{id,t}$ represents the log of export volume of seller i in destination d in year t, $\boldsymbol{x_{id,t}}$ captures seller characteristics, and $D_{id,t-1}$ captures total demand for seller products in the previous period.[61] We use this specification to explore whether network externalities have any impact on export volumes of Belgian firms. When exploring this channel for export starters, we focus on the year of entry, which means our sample is a pooled cross-section of individual entry events at the firm-destination level. In this case, we employ firm and destination-year fixed effects. When studying the effect of networks for incumbent exporters, we can exploit variation within exporter-destination pairs over time and include firm-year and firm-destination FE.

Below we present two sets of results. First, Tables 14 and 15 show that neither export signals nor peer productivity have any impact on export volumes of starters and incumbents. Second, table 16 shows that these results hold when controlling for endogenous export signals and linkages or when using more stringent FE specifications.[62]

[61] We use log employment, wages, and productivity to control for firm heterogeneity. Firm demand follows the definition outlined in section 3.3.

[62] Adding firm-destination FE in column 6 of table 16 controls for time-invariant firm or product appeal in the foreign market. Firm-year FE absorb any variation at the firm-year level which includes productivity spillovers from network peers and the selection correction term.

Table 14: Intensive margin – export starters

Signal type:	matching	non-matching	total	EEA	border	history
export signal	-0.0006	-0.0001	-0.0001	-0.0001	-0.0070	-0.0006
	(0.0081)	(0.0011)	(0.0011)	(0.0018)	(0.0045)	(0.0012)
peer TFP	0.0144	0.0143	0.0143	0.0144	0.0139	0.0141
	(0.0253)	(0.0253)	(0.0253)	(0.0253)	(0.0253)	(0.0253)
Peer characteristics	✓	✓	✓	✓	✓	✓
Destination experience	✓	✓	✓	✓	✓	✓
Firm characteristics	✓	✓	✓	✓	✓	✓
firm FE	✓	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓	✓
R^2	0.4149	0.4149	0.4149	0.4149	0.4149	0.4149
Observations	94,318	94,318	94,318	94,318	94,318	94,318

This table shows the impact of network effects on export volumes of export starters in the year of entry. Results are based on equation 13. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

Table 15: Intensive margin – incumbent exporters

Signal type:	matching	non-matching	total	EEA	border	history
export signal	0.0117	-0.0007	-0.0007	-0.0016	0.0020	0.0011
	(0.0201)	(0.0005)	(0.0005)	(0.0010)	(0.0061)	(0.0014)
peer TFP	-0.0096	-0.0115	-0.0114	-0.0113	-0.0100	-0.0095
	(0.0183)	(0.0183)	(0.0183)	(0.0183)	(0.0182)	(0.0183)
Peer characteristics	✓	✓	✓	✓	✓	✓
Destination experience	✓	✓	✓	✓	✓	✓
Firm characteristics	✓	✓	✓	✓	✓	✓
firm FE	✓	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓	✓
R^2	0.5314	0.5314	0.5314	0.5314	0.5314	0.5314
Observations	143,107	143,107	143,107	143,107	143,107	143,107

This table shows the impact of network effects on export volumes of incumbent exporters. Results are based on equation 13. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

Table 16: Network effects and the margins of trade – details

	extensive margin		intensive margin			
	(1)	(2)	(3)	(4)	(5)	(6)
matching signals	0.0043***	0.0092*	-0.0006	0.0134	0.0117	0.1322
	(0.0011)	(0.0052)	(0.0084)	(0.0686)	(0.0201)	(0.0996)
log employment	0.0476***	0.0477***	0.0178	0.0178	0.1700***	
	(0.0041)	(0.0042)	(0.0253)	(0.0267)	(0.0488)	
log wage	0.0127**	0.0122**	-0.0267	-0.0286	-0.1511***	
	(0.0060)	(0.0062)	(0.0380)	(0.0394)	(0.0504)	
log TFP	0.0325***	0.0306***	0.1846***	0.1762***	0.2884***	
	(0.0040)	(0.0041)	(0.0249)	(0.0260)	(0.0520)	
log export demand	0.0179***	0.0179***	0.1225***	0.1213***	0.3932***	0.0239***
	(0.0003)	(0.0003)	(0.0033)	(0.0034)	(0.0099)	(0.0083)
border dummy	0.0693***	0.0690***	0.1016***	0.1008***		
	(0.0024)	(0.0024)	(0.0162)	(0.0166)		
history dummy	-0.0140***	-0.0143***	0.0841***	0.0830***		
	(0.0024)	(0.0024)	(0.0201)	(0.0206)		
export propensity	0.1702***	0.1663***	1.061***	1.023***		
	(0.0115)	(0.0119)	(0.0635)	(0.0657)		
import dummy	0.0980***	0.0961***	-0.0633**	-0.0625**		
	(0.0034)	(0.0034)	(0.0263)	(0.0271)		
log peer size	-0.0005	2.9×10^{-5}	0.0070	0.0059	0.0127	
	(0.0019)	(0.0020)	(0.0132)	(0.0140)	(0.0157)	
peer TFP	0.0004	0.0010	0.0144	0.0185	-0.0096	
	(0.0027)	(0.0028)	(0.0200)	(0.0214)	(0.0183)	
IV		✓		✓		✓
selection correction		✓		✓		✓
firm FE	✓	✓	✓	✓	✓	
destination-year FE	✓	✓	✓	✓	✓	✓
firm-destination FE						✓
firm-year FE						✓
R^2	0.093	0.093	0.415	0.412	0.531	0.843
Observations	469,770	453,532	94,318	90,445	143,107	226,877

This table shows how export signals affect the entry probability and export volume of export starters and incumbents. Extensive margin regressions follow benchmark equation 7. Intensive margin regressions follow equation 13. Column pairs 1-2 and 3-4 show the impact of signals on entry probability and export volume of export starters. Column pair 5-6 shows how export signals affect export volumes of incumbent exporters. Columns 2, 4, and 6 control for endogenous export signals and linkages via instruments, selection correction, and fixed effects as explained in section 2.4. Note that column 6 uses firm-year FE to absorb any unobserved elements at the firm level, including selection bias from endogenous linkages. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.4 Demand uncertainty

In this section, we test whether export signals affect firm survival in foreign markets. To do so, we track the export status of entrants during the first n years after an entry event in period t. We estimate the following equation via LPM-FE

$$\text{exit}_{id,t+n} = \boldsymbol{a'x_{id,t+n}} + \rho D_{id,t+n-1} + \beta \sum_j s_{ij,t-1} y_{jd,t-1} + FE_i + FE_{d,t} + \varepsilon_{id,t+n} \tag{14}$$

where $\text{exit}_{id,t+n}$ is an indicator variable that is equal to 1 if seller i has exited market d in period $t+n$, $\boldsymbol{x_{id,t+n}}$ captures seller characteristics, and $D_{id,t+n-1}$ represents lagged demand for seller products in the previous period.[63]

The timing of export signals and outcomes differs from previous settings. What we are testing here is whether signals received *prior* to an entry in period t have any impact on exit decisions during the n^{th} period *after* entry. For $n = 1$, our regression sample consists of a cross section of exit decisions during the first period after entry. In this case, we can only include firm- and destination-year fixed effects to control for unobserved factors. For periods $t+2$ and $t+3$, we additionally add firm-year FE. To account for the impact of export experience before entry, we present separate results for the subset of sellers whose entry in destination d marks the first export experience at the firm level. Inexperienced firms might show a stronger response to a signal-induced reduction in demand uncertainty (Fernandes and Tang, 2014).

Table 17 summarizes the results across multiple time periods and specifications. Columns 1-3 show results for all export starters, whereas columns 4-6 focus on the subset of first-time exporters. Across all specifications, we do not find any impact of matching signals on exit behavior of Belgian firms.

[63] We use log employment, wages, and productivity to control for firm heterogeneity. Firm demand follows the definition outlined in section 3.3.

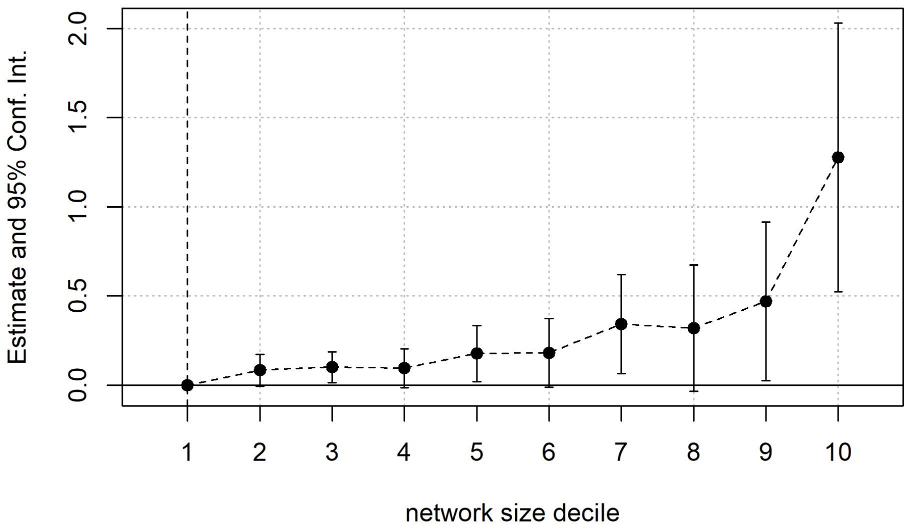
Table 17: Exit rates of export starters

	all starters			first-time exporters		
	t+1	t+2	t+3	t+1	t+2	t+3
matching signals	0.0006	-0.0005	0.0027	-0.0413	0.0129	0.0207
	(0.0026)	(0.0036)	(0.0036)	(0.0302)	(0.0305)	(0.0477)
Firm + destination controls	✓	✓	✓	✓	✓	✓
R^2	0.391	0.459	0.499	0.761	0.690	0.728
Observations	54,196	17,665	14,582	2,792	651	478
firm FE	✓	✓	✓	✓	✓	✓
destination-year FE	✓	✓	✓	✓	✓	✓

This table shows how export signals received in t-1 affect the exit probability of starters in the years after entry. Results are based on equation 14. Columns 1 to 3 study exit rates of all starters while columns 4 to 6 focus on firms without any export activity across destinations for two consecutive years prior to entry. Individual columns refer to different periods after entry. Period t+2, for example, studies the behavior of starters two years after entry. Standard errors in parentheses are clustered at the firm level. Significance codes: ***: 0.01, **: 0.05, *: 0.1.

D.5 Signal clarity

Figure 17: Signal clarity results by network size



Note: This figure plots matching signal coefficients β obtained from estimating equation 7 under a signal clarity specification. Coefficients are expressed relative to firms with the smallest networks and indicate how an increase in signal clarity affects the entry probability of Belgian firms. Further estimation details are described in section 4.4.